

STIC-Biotech/ChemLib

CRFE

155312 6/7 06h +02p

From: Epps-Ford, Janet  
Sent: Thursday, June 02, 2005 5:21 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/155,676

76  
570

6/8 8 (40 hrs)

Please search

SEQ ID NOs: 1, 3, 4, and 6 in all pending and commercial nucleic acid sequence databases. Reverse translation search in all amino acid pending and commercial databases.

SEQ ID NO: 2, 5, 7 in all pending and commercial amino acid databases.

Thanks,

Janet L. Epps-Ford, Ph.D.

Art Unit 1635

Mailbox: Remsen 2C18

Office: Remsen 2C05

Phone: 571-272-0757

Fax: 571-273-0757

1 na 1906  
3 na 2631  
4 na 1253  
6 na 4596  
2 aa 604  
5 aa 417  
7 aa 947

RECEIVED  
JUN - 3 2005  
(STIC)

me

STAFF USE ONLY

Searcher: Port  
Searcher Phone: 2-  
Date Searcher Picked up: 6/7  
Date Completed: 6/18  
Searcher Prep/Rev. Time: 20  
Online Time: 12

Type of Search

NA#: 4 AA#: 3  
Interference: ✓ SPDI: ✓  
S/L: ✓ Oligomer: ✓  
Encode/Transl: ✓  
Structure#: ✓ Text: ✓  
Inventor: ✓ Litigation: ✓

Vendors and cost where applicable

STN: ✓  
DIALOG: ✓  
QUESTEL/ORBIS: ✓  
LEXIS/NEXIS: ✓  
SEQUENCE SYSTEM: ✓  
WWW/Internet: ✓  
Other(Specify): ✓

**THIS PAGE BLANK (USPTO)**





# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number 155312

**TO: Janet Epps-Ford**  
**Location: rem/2C05/2C18**  
**Art Unit: 1635**  
**Tuesday, June 14, 2005**

**Case Serial Number: 09/155676**

**From: Toby Port**

**Location: Biotech-Chem Library**  
**REM1-A59**

**Phone: 272-2523**

**toby.port@uspto.gov**

### Search Notes

Dear Examiner Epps-Ford,

Here are the results of your search.

Please feel free to contact me if you have any questions.

Toby Port

**THIS PAGE BLANK (USPTO)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact ***the searcher or contact:***

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop-off or send completed forms to STIC Biotech-Chem Library, Remsen Bldg.



**"HIS PAGE BLANK (USPTO)"**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 647.812 Seconds  
(without alignments)  
17417.120 Million cell updates/sec

Title: US-09-155-676B-1  
Perfect score: 1906  
Sequence: 1 cattgggtcacgcgtggcg.....tcattccttgtaaaaaaaa 1906

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1899	99.6	1906	2 AAV03323	Aav03323 5' end of
2	1742	91.4	1782	4 AAD04333	Aad04333 Human TNF
3	1555	81.6	2873	4 AAD04335	Aad04335 Human TNF
4	1555	81.6	3139	4 AAD04334	Aad04334 Human TNF
5	1138.2	59.7	2248	4 AAH15665	Aah15665 Human CDN
6	792.8	41.6	6045	4 AAH18235	Aah18235 Human CDN
7	672.2	35.3	730	4 AAH04973	Aah04973 Human CDN
8	352.2	18.5	33147	4 AAK67282	Aak67282 Human imm
9	274.6	14.4	454	5 AAS64410	Aas64410 DNA encod
10	263.8	13.8	280	6 ABS72233	Abs72233 Human gen
11	215.8	11.3	542	12 ACH75347	Ach75347 Human gen
12	212.8	11.2	284	12 ACH89047	Ach89047 Human gen
13	195.6	10.3	463	5 AAS92200	Aas92200 DNA encod
14	135.4	7.1	474	5 AAS73549	Aas73549 DNA encod
15	133.4	7.0	190	3 AAC11927	Aac11927 Human sec
16	128	6.7	500	12 ACH75141	Ach75141 Human gen
17	124	6.5	124	12 ACH88841	Ach88841 Human gen
18	119	6.2	911	5 AAS64411	Aas64411 DNA encod
19	83.6	4.4	100	3 AAC12137	Aac12137 Human sec
20	66.8	3.5	1796	4 ABL09365	Ab109365 Drosophil

C 21	66.8	3.5	3796	4 ABL09364	Ab109364 Drosophil
C 22	63.8	3.3	3931	4 ABA09172	Aba09172 Human BG3
C 23	61.6	3.2	1253	2 AAV03325	Aav03325 Sequence
C 24	61	3.2	2000	8 ADA71938	Ada71938 Rice gene
C 25	60	3.1	214520	10 ADL13471	Adl13471 Osteoarth
C 26	59.6	3.1	263	6 ABL86718	Ab186718 Human ova
C 27	59.4	3.1	5884	3 AAC59953	Aac59953 Human sec
C 28	59.4	3.1	5884	8 ABZ73290	Abz73290 Secreted
C 29	59.4	3.1	5684	10 ABZ66903	Abz66903 Human sec
C 30	59	3.1	3927	10 ADF81729	Adf81729 Leukaemia
C 31	58.4	3.1	595	13 ADQ53383	Adq53383 Novel can
C 32	56.6	3.0	729	10 ADE07225	Ade07225 Novel cod
C 33	54.6	2.9	371	9 ACH37072	Ach37072 Human end
C 34	54.6	2.9	446	6 ABN95020	Abn95020 Gene #151
C 35	54.6	2.9	769	4 AA195747	Aa195747 Human neu
C 36	54.4	2.9	813	5 AAS85362	Aas85362 DNA encod
C 37	54	2.8	464	9 ACH27965	Ach27965 Human adu
C 38	53.2	2.8	476	7 ADS31085	Ads31085 Human gen
C 39	53	2.8	143	4 AAD04332	Aad04332 Human ire
C 40	52.2	2.7	941	5 ABA21208	Aba21208 Human ner
C 41	52.2	2.7	941	5 ABA21210	Aba21210 Human ner
C 42	52.2	2.7	961	12 ADL12492	Adl12492 Human ste
C 43	52.2	2.7	3916	13 ADR07577	Adr07577 Full leng
C 44	51.8	2.7	617	5 ABA14117	Ab14117 Human ner
C 45	50.4	2.6	7093	6 ABZ35416	Abz35416 Human gen

ALIGNMENTS

RESULT 1  
AAV03323  
ID AAV03323 standard; cDNA; 1906 BP.  
AC AAV03323;  
XX  
DT 15-APR-1998 (first entry)  
DE 5' end of clone 9, which encodes a TRAF2 binding protein.  
KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;  
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;  
KW intracellular signalling activity; acute hepatitis;  
KW autoimmune-induced cell death; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 94..1906  
FT /\*tag= a  
FT /codon= 127-129, aa:Ser  
FT /note= "no ATG start or STOP codon given"  
XX  
PN WO9737016-A1.  
XX  
PD 09-OCT-1997.  
XX  
PF 01-APR-1997; 97WO-IL000117.  
XX  
PR 02-APR-1996; 96IL-00117800.  
PR 26-AUG-1996; 96IL-00119133.  
XX  
PA (YEDA ) YEDA RES & DEV CO LTD.  
XX  
PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;  
XX  
DR WPI; 1997-503101/46.  
DR P-PSDB; AAW42400.  
XX  
PT DNA encoding tumour necrosis factor receptor-associated factor binding  
PT molecule - used for modulation or mediation in cells of the activity of  
PT NF-KB.  
XX



QY 1801 TAGTCTCATTTGAGCTCCTCGATCCAGTCTTTCTGAAGCTGTGTTTCTCTCGACTTTT 1860  
Db |||||||  
QY 1801 TAGTCTCATTTGAGCTCCTCGATCCAGTCTTTCTGAAGCTGTGTTTCTCTCGACTTTT 1860  
Db |||||||  
QY 1861 CATGTATGTGAGCCAAATAATTCCTTTCATTCTTCAATTCCTGAAAAA 1906  
Db |||||||  
QY 1861 CATGTATGTGAGCCAAATAATTCCTTTCATTCTTCAATTCCTGAAAAA 1906  
Db |||||||

## RESULT 2

AAD04333

ID AAD04333 standard; cDNA; 1782 BP.

XX

AC AAD04333;

XX

XX

DT 04-JUL-2001 (first entry)

XX

DE Human TNF receptor-associated factor (TRAF2) binding protein, IREN cDNA.

XX

KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;

KW TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; immunosuppressive;

KW nuclear factor-kappaB; NF-kappaB; cytostatic; tumour; AIDS;

KW acquired immune deficiency syndrome; rheumatic disease; apoptosis;

KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;

KW anorexia; anti-HIV; therapy; ss.

XX

OS Homo sapiens.

XX

XX

FH Key

CDS

FT 1.1626

FT /\*tag= a

FT /product= "Human TRAF2-binding protein, IREN"

FT

XX

XX

PN WO200116314-A1.

XX

XX

PD 08-MAR-2001.

XX

PF 31-AUG-2000; 2000WO-IL000517.

XX

PR 02-SEP-1999; 99IL-00131719.

XX

XX

PA (YEDA ) YEDA RES &amp; DEV CO LTD.

XX

XX

PI Wallach D, Malinin N, Sinha I, Leu S;

XX

DR WPI; 2001-281387/29.

DR

XX

XX

PT New DNA sequence encoding Tumor Necrosis Factor receptor associated

PT factor (TRAF) binding proteins (IREN) for treatment or prevention of

PT pathological conditions associated with NF-kappaB induction.

XX

XX

PS Claim 1; Fig 3B; 118pp; English.

XX

XX

CC The present cDNA sequence encodes human tumour necrosis factor (TNF)

CC receptor-associated factor (TRAF2) binding protein, IREN. A fragment of

CC this IREN molecule is capable of binding to TRAF2 protein at position 225

CC -501. The invention relates to human tumour necrosis factor (TNF)

CC extracts or transformed cell lines, in addition IREN can be used in  
CC diagnostic purposes for identifying disorders related to abnormal  
CC functioning of cellular effects mediated directly by TRAF proteins  
XX  
XX  
XX Sequence 1782 BP; 456 A; 458 C; 504 G; 364 T; 0 U; 0 Other;

Query Match 91.4%; Score 1742; DB 4; Length 1782;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1775; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 118 ATGAGCGGATCNCNGAACNATGACAAAAGACAAATTTCTGCTGAGCGACTGCTGGATGCA 177

Db |||||||

QY 178 GTGAAACAGTGCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCTCGATTCCGACAGC 237

Db |||||||

QY 238 AGGTCACCTGTCTGTGTGCCAGTTTGAAGCGTCTCTGAGCATGCTTGAAGAGAGT 297

Db |||||||

QY 121 AGGTCACCTGTCTGTGTGCCAGTTTGAAGCGTCTCTGAGCATGCTTGAAGAGAGT 180

Db |||||||

QY 298 CGAGGATTGCGCACTCACAGCGGCGAGATCAAGCAGCGAGCGGCTTTGCCAGCAAAACC 357

Db |||||||

QY 181 CGAGGATTGCGCACTCACAGCGGCGAGATCAAGCAGCGAGCGGCTTTGCCAGCAAAACC 240

Db |||||||

QY 358 GAAACAGAGCCCGTGTTCGTGCTACTACGTGAAGAGGTCTCTCAACAGCACGAGCTGCAG 417

Db |||||||

QY 241 GAAACAGAGCCCGTGTTCGTGCTACTACGTGAAGAGGTCTCTCAACAGCACGAGCTGCAG 300

Db |||||||

QY 418 CGCTTCTACTCTCCCTGCGCCACATCGCTCAGACGTGGCGGGGTCGCGCTGGCTGCGC 477

Db |||||||

QY 301 CGCTTCTACTCTCCCTGCGCCACATCGCTCAGACGTGGCGGGGTCGCGCTGGCTGCGC 360

Db |||||||

QY 478 TGTGCTCTCAACGAAACACTCCCTGGAGCGCTACCTGCACATGCTCTGCGCGACCGCTGC 537

Db |||||||

QY 361 TGTGCTCTCAACGAAACACTCCCTGGAGCGCTACCTGCACATGCTCTGCGCGACCGCTGC 420

Db |||||||

QY 538 AGGCTGAGCACTTTTATGAAGACTGGTCTTTGTGATGATGAAGAAAGTCCAGTATG 597

Db |||||||

QY 421 AGGCTGAGCACTTTTATGAAGACTGGTCTTTGTGATGATGAAGAAAGTCCAGTATG 480

Db |||||||

QY 598 CTTCTTACCATGGCAGCAGGTCTGAACCTCACTCTTTTGGGATTAACATCGAACAG 657

Db |||||||

QY 481 CTTCTTACCATGGCAGCAGGTCTGAACCTCACTCTTTTGGGATTAACATCGAACAG 540

Db |||||||

QY 658 GATTTGAACGGCGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAGAGTCAACG 717

Db |||||||

QY 541 GATTTGAACGGCGCAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAGAGTCAACG 600

Db |||||||

QY 718 CAGACGTGACCTCTCTGCTGAGAGGAGTCCACGCAAGGAGTGAAGAGTCTGTTGAGGAG 777

Db |||||||

QY 601 CAGACGTGACCTCTCTGCTGAGAGGAGTCCACGCAAGGAGTGAAGAGTCTGTTGAGGAG 660

Db |||||||

QY 778 ATCAGACCTCTCTGCTGCTGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCCTTG 836

Db |||||||

QY 661 ATCAGACCTCTCTGCTGCTGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCCTTG 720

Db |||||||

QY 837 CTTGCTGCTCAGGAGTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 896

Db |||||||

QY 721 CTTGCTGCTCAGGAGTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 780

Db |||||||

QY 897 AAGAAAGTCAACCAATATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGCGGAC 956

Db |||||||

QY 781 AAGAAAGTCAACCAATATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGCGGAC 840

Db |||||||

QY 957 GTGTTTAAAAAGACACCTCGGGCAGGGAGAGTCTAGAGACAACTCCGACCGCTCTCT 1016

Db |||||||

QY 841 GTGTTTAAAAAGACACCTCGGGCAGGGAGAGTCTAGAGACAACTCCGACCGCTCTCT 900

Db |||||||

QY 1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAA-TCAGAGC 1075

Db |||||||

QY 901 GTCAATATCATGTCCGCTTTGAAAGCCCTTCGGGCTTAACCTCAATGGAAAGTCAAGC 960

Db |||||||





Db 181 CGAGATTGGCACTCAGCGGCGAGCGATCAAGCAGCGACGCGGCTTTGGCAGCAAAACC 240  
Qy 358 GAAACAGAGCCGGTGTCTGTACTACGTGAAGAGGTCTCTCAACAGACGAGCTGCAG 417  
Db 241 GAAACAGAGCCGGTGTCTGTACTACGTGAAGAGGTCTCTCAACAGACGAGCTGCAG 300  
Qy 418 CGCTTCTACTCCCTGGCCCATGCGCTCAGACGTGGCCCGGGTGGCGCTGGCTGGCG 477  
Db 301 CGCTTCTACTCCCTGGCCCATGCGCTCAGACGTGGCCCGGGTGGCGCTGGCTGGCG 360  
Qy 478 TGTGCGCTCAACGAACACTCCCTGGAGCGCTACTCGACATGCTCTCTGGCCGAGCCGCTGC 537  
Db 361 TGTGCGCTCAACGAACACTCCCTGGAGCGCTACTCGACATGCTCTCTGGCCGAGCCGCTGC 420  
Qy 538 AGCTGAGCACTTTTATGAAGACTGTGTCTTTTGTGATGATGAAGAAAGTCCAGTATG 597  
Db 421 AGCTGAGCACTTTTATGAAGACTGTGTCTTTTGTGATGATGAAGAAAGTCCAGTATG 480  
Qy 598 CTTCTACCATGGCAGCAGTCTGAACCTCCTATCTCTTTCGGATTAAACATCGACAAG 657  
Db 481 CTTCTACCATGGCAGCAGTCTGAACCTCCTATCTCTTTCGGATTAAACATCGACAAG 540  
Qy 658 GATTGTAACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACG 717  
Db 541 GATTGTAACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACG 600  
Qy 718 CAGAACGTGACCTCTTGTCTGAAGAGTCCACCAAGAGTGAAGCAGCTGTTCAGGGAG 777  
Db 601 CAGAACGTGACCTCTTGTCTGAAGAGTCCACCAAGAGTGAAGCAGCTGTTCAGGGAG 660  
Qy 778 ATCAGAGCTCTCTGCGCTCTCATCTCATCAACCTGAAACAGAGACCGA-CGCTTG 836  
Db 661 ATCAGAGCTCTCTGCGCTCTCATCTCATCAACCTGAAACAGAGACCGA-CGCTTG 720  
Qy 837 CCGTCTGTGCCAGGATGTCAGTGTGATGCGCAATGCAAAAGGCGGAGAGAA 896  
Db 721 CCGTCTGTGCCAGGATGTCAGTGTGATGCGCAATGCAAAAGGCGGAGAGAA 780  
Qy 897 AAGAAAGTGAACCAATAATCTCATTTGATGATGAGGAAGTGAAGCAGCAACTCTGGGGAC 956  
Db 781 AAGAAAGTGAACCAATAATCTCATTTGATGATGAGGAAGTGAAGCAGCAACTCTGGGGAC 840  
Qy 957 GTGTTTAAAGACACCTGGGGCAGGGGAGGTCTAGAGACAACTCCCGACCGCTCTCT 1016  
Db 841 GTGTTTAAAGACACCTGGGGCAGGGGAGGTCTAGAGACAACTCCCGACCGCTCTCT 900  
Qy 1017 GTCAATATCATGTCCGCTTTGAAGCCCTTCCGGCCCTAACTCCCAATGGA-TCAGAGC 1075  
Db 901 GTCAATATCATGTCCGCTTTGAAGCCCTTCCGGCCCTAACTCCCAATGGAAGTCAAGC 960  
Qy 1076 AGCAACTCATGGAATAATTGATCCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT 1135  
Db 961 AGCAACTCATGGAATAATTGATCCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT 1020  
Qy 1136 GATGTGAAAGACATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGAAAC 1195  
Db 1021 GATGTGAAAGACATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGAAAC 1080  
Qy 1196 TCATCAGGAAGGACGACAGGGGCCACTCGAGTCCGCCGAGAACCACTGGAGGGAAC 1255  
Db 1081 TCATCAGGAAGGACGACAGGGGCCACTCGAGTCCGCCGAGAACCACTGGAGGGAAC 1140  
Qy 1256 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGTGTGACATGATCTCCGAC 1315  
Db 1141 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGTGTGACATGATCTCCGAC 1200  
Qy 1316 ATCTCTCTCCCTGTCAAGTGGGCTCTTACAGCCAGCAGATGCCCTCCGGAAGC 1375  
Db 1201 ATCTCTCTCCCTGTCAAGTGGGCTCTTACAGCCAGCAGATGCCCTCCGGAAGC 1260  
Qy 1376 CTGGAGACGGGACAGACAGAGGACCAAGTCTCCGATCTCCGATCTCCGATCACT 1435  
Db 1261 CTGGAGACGGGACAGACAGAGGACCAAGTCTCCGATCTCCGATCTCCGATCACT 1320

Qy 1436 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCCTGTGTA-CTTCTGCCTCA 1494  
Db 1321 GTGGAAGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCCTGTGTA-CTTCTGCCTCA 1380  
Qy 1495 GTGCCAGAGTCCATGACAAATTAGTGAACCTGCCGAGCCCACTGTGCCATGATGACAGG 1554  
Db 1381 GTGCCAGAGTCCATGACAAATTAGTGAACCTGCCGAGCCCACTGTGCCATGATGACAGG 1440  
Qy 1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTCCAGCGTGAATGGAG 1614  
Db 1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTCCAGCGTGAATGGAG 1500  
Qy 1615 CACTAGCCGCGCTCCGGCAAGAGTGGACACCTTGAAGAGGAGTGGTGAACAGGAG 1674  
Db 1501 CACTAGCCGCGCTCCGGCAAGAGTGGACACCTTGAAGAGGAGTGGTGAACAGGAG 1560  
Qy 1675 GAGCGCAGCGGATGAGGTCCAGGCGCTGGCCAG 1709  
Db 1561 GAGCGCAGCGGATGAGGTCCAGGCGCTGGCCAG 1595

## RESULT 4

AAD04334  
ID AAD04334 standard; cDNA; 3139 BP.

XX AAD04334;

XX AC AC (first entry)

XX 04-JUL-2001

XX Human TNF receptor-associated factor binding protein, IREN-10B cDNA.

XX Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;

XX TRAF2; TRAF2 binding protein; IkappaB Regulator; IREN-10B;

XX immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytosolic; tumour;

XX AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;

XX autoimmune disease; septic shock; graft-vs-host reaction; inflammation;

XX anorexia; anti-HIV; therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS l..2442

XX FT /\*tag= a

XX FT /product= "Human TRAF2-binding protein isoform, IREN-10B"

XX WO200116314-A1.

XX PD 08-MAR-2001.

XX 31-AUG-2000; 2000WO-IL000517.

XX 02-SEP-1999; 99IL-00131719.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX Wallach D, Malinin N, Sinha I, Leu S;

XX WPI; 2001-281387/29.

XX P-PSDB; AAE00684.

XX New DNA sequence encoding Tumor Necrosis Factor receptor associated

XX factor (TRAF) binding proteins (IREN) for treatment or prevention of

XX pathological conditions associated with NF-kappaB induction.

XX Claim 1; Fig 4; 118pp; English.

XX The present cDNA sequence encodes human tumour necrosis factor (TNF)

XX receptor-associated factor (TRAF2) binding protein isoform, IREN-10B. A

XX fragment of this IREN-10B molecule is capable of binding to TRAF2 protein

XX at position 225-501. The invention relates to human tumour necrosis

XX factor (TNF) receptor-associated factor (TRAF2) binding protein

XX designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and

CC their corresponding cDNA molecules. IREN is useful for  
CC modulating/mediating the activity of transcription factor NF (Nuclear  
CC Factor)-kappaB or any other intracellular signalling activity mediated by  
CC TRAF2. IREN is useful in the prevention and treatment of a pathological  
CC condition associated with NF-kappaB induction (abnormal) e.g. AIDS  
CC (acquired immune deficiency syndrome), autoimmune diseases, tumours,  
CC rheumatic diseases, anorexia, septic shock and graft-vs-host reactions.  
CC IREN also plays an important role in the control of inflammation and  
CC other non-apoptotic effects of TNF as well as in the control of  
CC apoptosis. The invention also relates to method for screening,  
CC identifying and producing a molecule capable of modulating activities  
CC mediated by IREN. IREN antibodies are useful for the purification of new  
CC proteins from different sources, including cell extracts or transformed  
CC cell lines, in addition IREN can be used in diagnostic purposes for  
CC identifying disorders related to abnormal functioning of cellular effects  
CC mediated directly by TRAF proteins  
XX

SQ Sequence 3139 BP; 797 A; 851 C; 875 G; 616 T; 0 U; 0 Other;

Query Match 81.6%; Score 1555; DB 4; Length 3139;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY	118	ATGAGCGGATCNCGAAACATACAAAAGACAATTTCTGTGGAGCGACTGCTGATGCA	177
DB	1	ATGAGCGGATCAGAAACAATGACAAAAGACAATTTCTGTGGAGCGACTGCTGATGCA	60
QY	178	GTGAACACAGTGCAGATCCGCTTTNGAGGAGAGAGAGATTGCCCTCGATTCCGACAGC	237
DB	61	GTGAACACAGTGCAGATCCGCTTTGAGGAGAGAGAGATTGCCCTCGATTCCGACAGC	120
QY	238	AGGTCACCTGTCGTGTGCCAGTTTGAAGCGCTCTGCAGCATGGCTTTGAAGAGAGT	297
DB	121	AGGTCACCTGTCGTGTGCCAGTTTGAAGCGCTCTGCAGCATGGCTTTGAAGAGAGT	180
QY	298	CSAGGATTGGCACTCACAGCGCGCAGCATCAAGCAGCAGCGGGCTTTGCCAGCAAAACC	357
DB	181	CSAGGATTGGCACTCACAGCGCGCAGCATCAAGCAGCAGCGGGCTTTGCCAGCAAAACC	240
QY	358	GAACAGAGCCGCTGTTCTGGTACTACGTGAAGAGAGTCTCAACAGCAGCAGCTGCAG	417
DB	241	GAACAGAGCCGCTGTTCTGGTACTACGTGAAGAGAGTCTCAACAGCAGCAGCTGCAG	300
QY	418	CGCTTCTACTCCCTCGCGCCACATCGCTCAGACGTGGCCGGGGTGGCGCTGCTGCGC	477
DB	301	CGCTTCTACTCCCTCGCGCCACATCGCTCAGACGTGGCGCGGGTGGCGCTGCTGCGC	360
QY	478	TGTGCCCTCAACGAACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCGCAGCGCTGC	537
DB	361	TGTGCCCTCAACGAACTCCCTGGAGCGCTACCTGCACATGCTCCTGGCGCAGCGCTGC	420
QY	538	AGGCTGAGCATTTTATGAAGACTGGTCTTTTGTGATGAAGAAAGTCCAGTATG	597
DB	421	AGGCTGAGCATTTTATGAAGACTGGTCTTTTGTGATGAAGAAAGTCCAGTATG	480
QY	598	CTTCCTACCATGGCAGCAGGTCTGAATCCATCTCTTTGGGATTAACATCGACAAACAG	657
DB	481	CTTCCTACCATGGCAGCAGGTCTGAATCCATCTCTTTGGGATTAACATCGACAAACAG	540
QY	658	GATTGGAACGGGAGAGTAAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACG	717
DB	541	GATTGGAACGGGAGAGTAAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACG	600
QY	718	CAGAACGTGACCTCTTGTGAAGAGTCCACGCAAGAGTGAGCAGCCTGTTTCAGGGAG	777
DB	601	CAGAACGTGACCTCTTGTGAAGAGTCCACGCAAGAGTGAGCAGCCTGTTTCAGGGAG	660
QY	778	ATCACAGCCTCTCTGCGCTCTCCATCTCATCAAACTGACAGGAGCGGA-CCCTTG	836
DB	661	ATCACAGCCTCTCTGCGCTCTCCATCTCATCAAACTGACAGGAGACCGACCCCTTG	720
QY	837	CTGTCTGTGCCAGGAATGTCACTGCTGATGCCAAATGCAAAAGAGCGGAGAGAGAA	896
DB			

DB	721	CCTGCTGTCCAGGAATGTCACTGCTGATGCCAAATGCAAAAGGAGCGGAGAGAA	780
QY	897	AAGAAAGTGACCAACAATAATCTCAATTTGATGATGAGAAAGATGAGCAGAACTCTGGGAC	956
DB	781	AAGAAAGTGACCAACAATAATCTCAATTTGATGATGAGAAAGATGAGCAGAACTCTGGGAC	840
QY	957	GTGTTTAAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCT	1016
DB	841	GTGTTTAAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCT	900
QY	1017	GTCAATATCATGTCCGCTTTTGAAGAGCCCTTCCGGGCTTAACCTCAATGGAA-TCAGAGC	1075
DB	901	GTCAATATCATGTCCGCTTTTGAAGAGCCCTTCCGGGCTTAACCTCAATGGAAAGTCAGAGC	960
QY	1076	AGCAACTCATGGAATAATGATTCCTGTCTTTGAACCGGGAGTTTGGGTACCAAGAGCTT	1135
DB	961	AGCAACTCATGGAATAATGATTCCTGTCTTTGAACCGGGAGTTTGGGTACCAAGAGCTT	1020
QY	1136	GATGTGAAAGATCGATGATGAAGATGTGATGAAACCGAAAGATGACGTGTATGGAAAC	1195
DB	1021	GATGTGAAAGATCGATGATGAAGATGTGATGAAACCGAAAGATGACGTGTATGGAAAC	1080
QY	1196	TCATCAGAAAGGAAGACACAGGGCCACTCGAGGTGCGCCGAGAGCCACTGGAAGGAAAC	1255
DB	1081	TCATCAGAAAGGAAGACACAGGGCCACTCGAGGTGCGCCGAGAGCCACTGGAAGGAAAC	1140
QY	1256	ACCTGCCCTCTCCAGATGCAAGCTGGGCTCCGCTGAAGGTGCTGCAATGACTCCGAC	1315
DB	1141	ACCTGCCCTCTCCAGATGCAAGCTGGGCTCCGCTGAAGGTGCTGCAATGACTCCGAC	1200
QY	1316	ATCCTCTTCCCTGTCACTGTCGGCTCCCTACAGCCAGCAGATGCCCCCTCGGAAGC	1375
DB	1201	ATCCTCTTCCCTGTCACTGTCGGCTCCCTACAGCCAGCAGATGCCCCCTCGGAAGC	1260
QY	1376	CTGAGAAACCGGACAGGACACAGAGCACAGTTCTCCCGATCCTCGACTTCGATACAGT	1435
DB	1261	CTGAGAAACCGGACAGGACACAGAGCACAGTTCTCCCGATCCTCGACTTCGATACAGT	1320
QY	1436	GTGGAAGCCAGCTCTCCAGGCCAGGAAGTCCCTGAGCAGCGCTGTA-CTTCTGCTCA	1494
DB	1321	GTGGAAGCCAGCTCTCCAGGCCAGGAAGTCCCTGAGCAGCGCTGTA-CTTCTGCTCA	1380
QY	1495	GTGCCAGAGTCCATGACAAATTAGTAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG	1554
DB	1381	GTGCCAGAGTCCATGACAAATTAGTAACTGCGCCAGGCCACTGTGGCCATGATGAACAGG	1440
QY	1555	AAGGATGAGCTGAGGAGGAGAAACAGATCACTGCGAAACCTGCTCAGCGTGAAGAG	1614
DB	1441	AAGGATGAGCTGAGGAGGAGAAACAGATCACTGCGAAACCTGCTCAGCGTGAAGAG	1500
QY	1615	CATCTCAGCCGCTCCCGCAAGAGTGGACACCTTCAAAAGGAGGTGGCTGAACAGGAG	1674
DB	1501	CATCTCAGCCGCTCCCGCAAGAGTGGACACCTTCAAAAGGAGGTGGCTGAACAGGAG	1560
QY	1675	GAGCGCAGGGCATGAAGGTCCAGGCGCTGGCCAG	1709
DB	1561	GAGCGCAGGGCATGAAGGTCCAGGCGCTGGCCAG	1595

## RESULT 5

AAH15665

ID AAH15665 standard; cDNA; 2248 BP.

XX AAH15665;

XX AC  
XX AC  
DT 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:14016.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX

PN EP1074617-A2.  
 XX 07-FEB-2001.  
 XX 28-JUL-2000; 2000EP-00116126.  
 XX 29-JUL-1999; 99JP-00248036.  
 XX 27-AUG-1999; 99JP-00300253.  
 PR 11-JAN-2000; 2000JP-00118776.  
 PR 02-MAY-2000; 2000JP-00183767.  
 PR 09-JUN-2000; 2000JP-00241899.  
 XX (HELI-) HELIX RES INST.  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
 PT length cDNAs defined in the specification, and for the detection and/or  
 PT diagnosis of the abnormality of the proteins encoded by the full-length  
 PT cDNAs.  
 XX Claim 8; SEQ ID NO 14016; 2537pp + Sequence Listing; English.  
 XX The present invention describes primer sets for synthesizing 5602 full-  
 CC length cDNAs defined in the specification. Where a primer set comprises:  
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
 CC complementary strand of a polynucleotide which comprises one of the 5602  
 CC nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence and 3'-end sequence is selected from those defined in the  
 CC specification. The primer sets can be used in antisense therapy and in  
 CC gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
 CC oligonucleotides, all of which are used in the exemplification of the  
 CC present invention  
 XX Sequence 2248 BP; 548 A; 548 C; 618 G; 534 T; 0 U; 0 Other;  
 Query Match 59.7%; Score 1138.2; DB 4; Length 2248;  
 Best Local Similarity 94.7%; Pred. No. 1.6e-299;  
 Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;  
 77 GCCGGCGCGCGCGCANGCACCAGCGCCCGGGGAGAGCCATGAGCGGATCNCNGAACN 136  
 22 GCGCGCGCGCGCGCAGGACCGCGCCCGGGGAGAGCCATGAGCGGATCAGAGA 81  
 137 ATGACAAAGACAAATTTCTCTGAGGCGACTGCTGGATGAGTGAACAGTGCAGATCC 196  
 82 ATGACAAAGACAAATTTCTCTGAGGCGACTGCTGGATGAGTGAACAGTGCAGATCC 141  
 197 GCTTTNGAGGAGAAAGAGATGCTCGGATTTCCGACAGCAGGCTCACCTGTCGTGTG 256  
 142 GCTTTGAGGAGAAAGAGATGCTCGGATTTCCGACAGCAGGCTCACCTGTCGTGTG 201  
 257 CCCAGTTTGAAGCGCTCCTGCAGCATGGCTTGAAGAGAGTTCGAGGATTCGCACTCACAG 316  
 202 CCCAGTTTGAAGCGCTCCTGCAGCATGGCTTGAAGAGAGTTCGAGGATTCGCACTCACAG 261  
 317 CGGAGCGATCAAGAGGCGAGCGGCTTTCCAGCAAAACCGAAACAGAGCCCGGTGTCT 376

Db 262 CGGAGGATCAAGAGGAGCGGGCTTTGCCAGGAAACCGAAACAGAGCCCGTGTCT 321  
 Qy 377 GGTACTACGTGAAGAGGTCTCTCAACAAAGCAGAGCTGACGCGCTTCTACTCCCTGCGCC 436  
 Db 322 GGTACTAGGTGAAGAGGTCTCTCAACAAAGCAGAGCTGACGCGCTTCTACTCCCTGCGCC 381  
 Qy 437 ACATCGCTCAGAGCTGGGCGGGGTGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 496  
 Db 382 ACATCGCTCAGAGCTGGGCGGGGTGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 441  
 Qy 497 CCCTGGAGGCTACCTGTCATGCTCTCTGGCCGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 556  
 Db 442 CCCTGGAGGCTACCTGTCATGCTCTCTGGCCGACGCTGACGCTGACGCTGACGCTGACGCTGACGCT 501  
 Qy 557 AAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGATATGCTTCTTCCATGCGCAGCAG 616  
 Db 502 AAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGATATGCTTCTTCCATGCGCAGCAG 561  
 Qy 617 GTCTGAATCCATCTCTTTGGCATTAACATCAACAAAGGATTTGAACGGGCGAGATA 676  
 Db 562 GTCTGAATCCATCTCTTTGGCATTAACATCAACAAAGGATTTGAACGGGCGAGATA 621  
 Qy 677 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAGAGCTGACCTCTCTTGC 736  
 Db 622 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAGAGCTGACCTCTCTTGC 681  
 Qy 737 TGAAGAGTCCACGCAAGGAGTGAAGAGTCTTTCAGGGAGATCAACAGCTCTCTCTGCCG 796  
 Db 682 TGAAGAGTCCACGCAAGGAGTGAAGAGTCTTTCAGGGAGATCAACAGCTCTCTCTGCCG 741  
 Qy 797 TCTCATCTCTCAAAACCTGAAACAGAGAGCCGA-CCCTTGGCTGTGTGTTCAGGAAATG 855  
 Db 742 TCTCATCTCTCAAAACCTGAAACAGAGAGCCGA-CCCTTGGCTGTGTGTTCAGGAAATG 801  
 Qy 856 TCAGTCTGATGCTCAAAATGTCARAAAGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915  
 Db 802 TCAGTCTGATGCTCAAAATGTCARAAAGGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861  
 Qy 916 TCTCATTTGATGATGAGGAGATGAGCAGAACTCTCGGGACGTGTTTAAAGAGACACCTG 975  
 Db 862 TCTCATTTGATGATGAGGAGATGAGCAGAACTCTCGGGACGTGTTTAAAGAGACACCTG 921  
 Qy 976 GGGCAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 1035  
 Db 922 GGGCAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 981  
 Qy 1036 TTGAAAGCCCTTCCGGGCTTAACTCCAAATGGA-TCAGAGCAGCAACTCATGGAATG 1094  
 Db 982 TTGAAAGCCCTTCCGGGCTTAACTCCAAATGGAAGTCAAGAGCAGCAACTCATGGAATG 1041  
 Qy 1095 ATTTCCTGTCTTTGAACGGGAGTTCGGGTACCAAGAGCTTGTATGAAAGAGCATCGATG 1154  
 Db 1042 ATTTCCTGTCTTTGAACGGGAGTTCGGGTACCAAGAGCTTGTATGAAAGAGCATCGATG 1101  
 Qy 1155 ATGAAGATGTGATGAAACGAAAGATGACGTGTATGGAACCTCATCAAGAGAGAGAGAGAGAGAGAG 1214  
 Db 1102 ATGAAGATGTGATGAAACGAAAGATGACGTGTATGGAACCTCATCAAGAGAGAGAGAGAGAGAGAG 1161  
 Qy 1215 GGGGCCACTCGGAGTCCCGGAGAGCCACTGGAAGGAGACACTGCCTCTCCAGATGC 1274  
 Db 1162 GGGGCCACTCGGAGTCCCGGAGAGAGTAAAGTTTGTGTAAAGGTGAGTCTTCCAGTCCC 1221  
 Qy 1275 ACAGCTGGGCTCCGCTGAAGGCTGCTGCACAAATGACTCCGACATCTCTTCCCTGTCTAGTG 1334  
 Db 1222 CAAGGCTGAGTCCCGTGGCGTGTCTCAGCTCACTGCAACCTCCACCTCCCGGTTTTCAG 1281  
 Qy 1335 GC 1336  
 Db 1282 GC 1283





PR 14-AUG-2000; 2000US-0225270P.  
 PR 14-AUG-2000; 2000US-0225447P.  
 PR 14-AUG-2000; 2000US-0225757P.  
 PR 14-AUG-2000; 2000US-0225758P.  
 PR 14-AUG-2000; 2000US-0225759P.  
 PR 18-AUG-2000; 2000US-0226279P.  
 PR 22-AUG-2000; 2000US-0226681P.  
 PR 22-AUG-2000; 2000US-0226686P.  
 PR 22-AUG-2000; 2000US-0227182P.  
 PR 23-AUG-2000; 2000US-0227009P.  
 PR 30-AUG-2000; 2000US-0228924P.  
 PR 01-SEP-2000; 2000US-0229287P.  
 PR 01-SEP-2000; 2000US-0229343P.  
 PR 01-SEP-2000; 2000US-0229344P.  
 PR 01-SEP-2000; 2000US-0229345P.  
 PR 05-SEP-2000; 2000US-0229509P.  
 PR 05-SEP-2000; 2000US-0229513P.  
 PR 06-SEP-2000; 2000US-0230437P.  
 PR 06-SEP-2000; 2000US-0230438P.  
 PR 08-SEP-2000; 2000US-0231242P.  
 PR 08-SEP-2000; 2000US-0231243P.  
 PR 08-SEP-2000; 2000US-0231244P.  
 PR 08-SEP-2000; 2000US-0231413P.  
 PR 08-SEP-2000; 2000US-0231414P.  
 PR 08-SEP-2000; 2000US-0232080P.  
 PR 08-SEP-2000; 2000US-0232081P.  
 PR 12-SEP-2000; 2000US-0231968P.  
 PR 14-SEP-2000; 2000US-0232397P.  
 PR 14-SEP-2000; 2000US-0232398P.  
 PR 14-SEP-2000; 2000US-0232399P.  
 PR 14-SEP-2000; 2000US-0232400P.  
 PR 14-SEP-2000; 2000US-0232401P.  
 PR 14-SEP-2000; 2000US-0233063P.  
 PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 23-SEP-2000; 2000US-0234997P.  
 PR 23-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 13-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 20-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241221P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241820P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.

PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 11-DEC-2000; 2000US-0251990P.  
 PR 05-JAN-2001; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 WPI, 2001-483426/52.  
 Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 useful for preventing, diagnosing and/or treating cancers and metastasis.  
 Disclosure; SEQ ID NO 22094; 3071pp + Sequence Listing; English.  
 AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
 amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 activity, and can be used in gene therapy and vaccine production. (I)  
 proteins and polynucleotides may be used in the prevention, diagnosis and  
 treatment of diseases associated with inappropriate (I) expression. For  
 example, they may be used to treat disorders associated with decreased  
 expression by rectifying mutations or deletions in a patient's genome  
 that affect the activity of (I) by expressing inactive proteins or to  
 supplement the patients own production of (I). Additionally, (I)  
 polynucleotides may be used to produce the secreted (I), by inserting the  
 nucleic acids into a host cell and culturing the cell to express the  
 protein. (I) proteins and polynucleotides may be used to prevent,  
 diagnose and treat immune/haematopoietic-related diseases, especially  
 cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 to AAK87694 represent human immune/haematopoietic antigen genomic  
 sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 represent sequences used in the exemplification of the present invention  
 Query Match 18.5%; Score 352.2; DB 4; Length 33147;  
 Best Local Similarity 96.4%; Pred. No. 1.8e-84;  
 Matches 371; Conservative 0; Mismatches 13; Indels 1; Gaps 1;  
 SQ Sequence 33147 BP; 8370 A; 7682 C; 8382 G; 8713 T; 0 U; 0 Other;

Qy	862	CTGATGCCAAATGCAAAAAGAGCGGGAAGAAAAAGAAAGTCACCAACATAATCTCAT	921
Db	8894	CAGATGCCAAATGCAAAAAGAGCGGGAAGAAAAAGAAAGTCACCAACATAATCTCAT	8953
Qy	922	TTGATGATGAGGAAGATGAGCAGAACTCTCGGGACGTTGTAAAGACACCTCGGGCAG	981
Db	8954	TTGATGATGAGGAAGATGAGCAGAACTCTCGGGACATGTTAAAGACACCTCGGGCAG	9013
Qy	982	GGGAGAGCTCAGAGGACAATCCGACCGCTCTCTGTCAATATCATGTGCGCCTTTGAAA	1041
Db	9014	GGGAGAGCTCAGAGGACAATCCGACCGCTCTCTGTCAATATCATGTGCGCCTTTGAAA	9073
Qy	1042	GCCCTTTGGGCGCTTAATCCAATGGAA-TCAGACGACCACTCATGAAAAATTGATTCCT	1100
Db	9074	GCCCTTTGGGCGCAATCCAATGGAAGTCAGACGACCACTCGTGAAAAATTGATTCCT	9133
Qy	1101	TGCTTTGAACGGGGAGTTTGGGTACAGAAAGCTTGATGTGAAAGCATCGATGATGAAG	1160
Db	9134	TGCTTTGAACAGGGAGTTTGGGTACAGAAAGCTTGATGTGAAAGCATCGATGATGAAG	9193
Qy	1161	ATGTGGATGAAACGGAAGATGACGTGTATGGAACCTCATCAGGAAGGAAGCACAGGGGCC	1220
Db	9194	ATGTGGATGAAACGGAAGATGACGTGTATGGAACCTCATCAGGACGGAAGCACAGGGGCC	9253
Qy	1221	ACTCGGAGTCGCCCGAAGCCACT	1245
Db	9254	ACTCAGAGTCGCCCGAAGTAACT	9278

## RESULT 9

AAS64410

ID AAS64410 standard: cDNA: 454 BP.

XX

AC AAS64410;

XX  
DT 13-FEB-2002 (first entry)

XX  
DE DNA encoding novel human diagnostic protein #214.  
XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

**⊗**

PF 30-MAR-2001; 2001WO-US008631.

XX  
21 MAY 2000  
2000

PR 31-MAR-2000; 2000US-  
DB 22-MAR-2000; 2000US-

PK	
XX	(HYSE-) HYSEQ INC.
XX	23-AUG-2000; 2000US-00649167.
PA	
XX	
PI	Dmanac RT, Liu C, Tang YT;
XX	
DR	WPI: 2001-639362/73.
DR	P-PSDB; ABG00223.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published](http://wipo.int/pub/published) pct sequences

XX

```
Query Match      14.4%; Score 274.6; DB 5; Length 454;
Best Local Similarity 96.2%; Pred. No. 3.3e-64;
Matches 277: Conservative 0; Mismatches 11; Indels 0
```

Qy		77	GCCGGCGGCGCGGCAGNACGCCGGGCGGGGNNAGCNCCTATGACGGGATCNCGAA	CN	136
Db		27	GGCGGGCGGCGGTGCAGGCACCGTTTCGGGGAGAGGCACCATTAGCGGATCACAGA	CA	86
Qy		137	ATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCAGTGAACAGTGCCAGATCC		196
Db		87	ATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCAGTGAACAGTGCCAGATCC		146
Qy		197	GCTTTNGAGGGAGAAAGGAGATTGCTCGGATTCGACAGCAGGGTCACTGTCTGTGTG		256
Db		147	GCTTTGGAGGGAGAAAGGAGATTGCTCGGATTCGACAGCAGGGTCACTGTCTGTGTG		206
Qy		257	CCCAGTTTGAGCCGTCCTGCAGCATGGCTTGAAAGAGGAGTCCGAGATTGGCACTCACAG		316
Db		207	CCCAGTTTGAGCCGTCCTGCAGCATGGCTTGAAAGAGGAGTCCGAGATTGGCACTCACAG		266
Qy		317	CGGACGGATCAAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAAA	CAG	364
Db		267	CGGACGGATCAAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAAA	CAG	314

## RESULT 10

RECEIVED TO  
ABS72233

ID ABS72233 standard: cDNA: 280 bp.

2000

AC ABS72233;

XXXX

DT 03-DEC-2002 (first entry)

Human gene trapped sequence (GTS) #193.

Human; gene trapped sequence; GTS; gene; ss; cancer; autoimmune disease; lupus; scleroderma; Crohn's disease; multiple sclerosis; immune disorder; inflammatory bowel disease; schizophrenia; psychosis; osteoarthritis; inflammatory disorder; diabetes; skin disorder; acne; eczema; asthma; rheumatoid arthritis; hypertension; atherosclerosis; Alzheimer's disease; cardiovascular disease; Parkinson's disease; osteoporosis; infertility; viral infection; parasitic infection; fungal infection; bacterial infection; forensic analysis; cellular differentiation.

OS Homo sapiens.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

PN US2002095031

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

PD 18-JUL-2002.

XXXX



PA (NEHL/) NEHLS M C.  
PA (ZAMB/) ZAMBROWICZ B.  
PA (SAND/) SANDS A T.  
XX  
PI Nehls MC, Zambrowicz B, Sands AT;  
XX WPI; 2002-656030/70.  
XX  
XX New isolated or purified human gene trapped sequences, useful for gene  
PT discovery, as markers for gene expression analysis, identifying and  
PT mapping the coding regions of human genome, or determining the genetic  
PT basis of human disease.  
XX  
XX Claim 1; SEQ ID NO 201; 36pp; English.  
PS  
XX The invention relates to isolated or purified polynucleotides that  
CC correspond to human gene trapped sequences (GTSs). The human GTSs are  
CC useful for gene discovery and as markers for gene expression analysis,  
CC for identifying and mapping the coding regions of the mammalian,  
CC particularly human, genome, for forensic analysis, and for determining  
CC the genetic basis of human disease. The peptides and proteins encoded by  
CC the polynucleotides are useful for generating antibodies, as reagents in  
CC diagnostic assays and in identifying other cellular gene products  
CC involved in the regulation of development and cellular differentiation of  
CC various cell types. The peptides are also useful as reagents in assays  
CC for screening of compounds used in treating disorders affecting  
CC development and cell differentiation. The GTSs are also useful in  
CC treating or ameliorating diseases associated with the expression of  
CC mutant or normal variants of the GTSs, e.g. cancer, autoimmune diseases,  
CC lupus, scleroderma, Crohn's disease, multiple sclerosis, inflammatory  
CC bowel disease, immune disorders, schizophrenia, psychosis, inflammatory  
CC disorders, diabetes, skin disorders such as acne or eczema,  
CC osteoarthritis, rheumatoid arthritis, hypertension, atherosclerosis,  
CC cardiovascular diseases, Alzheimer's disease, Parkinson's disease,  
CC osteoporosis, asthma, infertility, and viral, parasitic, fungal or  
CC bacterial infections. This sequence represents a human GTS of the  
CC invention  
XX  
SQ Sequence 280 BP; 71 A; 66 C; 91 G; 51 T; 0 U; 1 Other;  
Query Match 13.8%; Score 263.8; DB 6; Length 280;  
Best Local Similarity 98.6%; Pred. No. 2.3e-61;  
Matches 276; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
QY 1464 GTCTCTGAGCAGCTGTGA-CTTCTGCTCAGTGCCAGAGTCCATGACAAATTAGTGAAC 1522  
Db 1 GGCCTATGAGCAGCCTGNTACCTTCTGCTCAGTGCCAGAGTCCATGACAAATTAGTGAAC 60  
QY 1523 TGGCCAGGCCACTGTGGCCATGATCAACAGGAAGGATGAGCTGGAGGAGGAGACAGAT 1582  
Db 61 TGGCCAGGCCACTGTGGCCATGATCAACAGGAAGGATGAGCTGGAGGAGGAGACAGAT 120  
QY 1583 CACTGCGAACTGTCTGACGGTGAGATGAGACATCTAGCCCGCTCCGGCAAGAGGTGG 1642  
Db 121 CACTGCGAACTGTCTGACGGTGAGATGAGACATCTAGCCCGCTCCGGCAAGAGGTGG 180  
QY 1643 ACACCTTGAAGGAGGTGGCTGACAGAGGCGGAGGCGGATGAGTCCAGGCGC 1702  
Db 181 ACACCTTGAAGGAGGTGGCTGACAGAGGCGGAGGCGGATGAGTCCAGGCGC 240  
QY 1703 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTTCTAAC 1742  
Db 241 TGGCCAGCTATCTTTGCTATTTTGTGAGGAGATTTCTAAC 280  
RESULT 11  
ID ACH75347/c  
XX ACH75347 standard; DNA; 542 BP.  
XX ACH75347;  
XX 29-JUL-2004 (first entry)  
XX

DE Human genome derived single exon probe #8542.  
XX Human; probe; ss; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
XX (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANK/) HANKEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
XX WPI; 2004-119264/12.  
XX  
XX New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
XX Claim 15; SEQ ID NO 8542; 80pp; English.  
PS  
XX The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX  
SQ Sequence 542 BP; 135 A; 143 C; 129 G; 135 T; 0 U; 0 Other;  
Query Match 11.3%; Score 215.8; DB 12; Length 542;  
Best Local Similarity 95.7%; Pred. No. 4e-48;  
Matches 244; Conservative 0; Mismatches 7; Indels 4; Gaps 2;  
QY 614 CAGGTCTGAATCCATCTCTTTGCGATTACATCGACACAGGATTGTAACGGGCAGA 673



Db 495 CAGTCCGAACTCATCTCTTTCCGATTAACATGCAACAAGGATTTGAACGGGCGAGA 436  
Qy 674 GTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAAGCGTGACCTCCT 733  
Db 435 GTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAAGTGGA---CCT 379  
Qy 734 TGCTGAAGGAGTGACCGAAGGAGTGAGCGCTGTTCAGGGAGATCACAGCTCTCTCTG 793  
Db 378 TGCTGAAGGAGTGACCGAAGGAGTGAGCGCTGTTCAGGGAGATCACAGCTCTCTCTG 319  
Qy 794 CCGTCTCCATCCCTCATCAACCTGAACAGGAGACCGA---CCCTTGCCTGTCTCTCCAGGA 852  
Db 318 CCATCTCCATCCCTCATCAACCTGAACAGGAGACCGACCTTGCCTGTCTCTCCAGGA 259  
Qy 853 ATGTCAGTGTGTG 867  
Db 258 ATGTCAGTGTGTG 244

RESULT 12  
ACH89047/c  
ID ACH89047 standard; DNA; 284 BP.  
AC ACH89047;  
XX  
XX 29-JUL-2004 (first entry)  
XX Human genome derived single exon probe #22242.  
XX Human; probe; ss; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX Homo sapiens.  
XX  
XX US2003194704-A1.  
XX 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX  
XX New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
XX Claim 1; SEQ ID NO 22242; 80pp; English.  
XX  
XX The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid

CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX  
SQ Sequence 284 BP; 60 A; 78 C; 74 G; 72 T; 0 U; 0 Other;

Query Match 11.2%; Score 212.8; DB 12; Length 284;  
Best Local Similarity 95.8%; Pred. No. 1.9e-47;  
Matches 241; Conservative 0; Mismatches 7; Indels 4; Gaps 2;  
Qy 617 GTCTGAACCTCCATCTCTTTGCGATTAAACATGCAACAAGGATTTGAACGGGCGAGTA 676  
Db 284 GTCCGAACCTCCATCTCTTTGCGATTAAACATGCAACAAGGATTTGAACGGGCGAGTA 225  
Qy 677 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAAGCGTGACCTCTCTG 736  
Db 224 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAAGTGCA---CCTTGC 168  
Qy 737 TGAAGGAGTCCACGCAAGGAGTGAGCGCTGTTCAGGGAGATCACAGCTCTCTCTCCG 796  
Db 167 TGAAGGAGTCCACGCAAGGAGTGAGCGCTGTTCAGGGAGATCACAGCTCTCTCTCCCA 108  
Qy 797 TCTCCATCTCTCATCAACCTGAACAGGAGACCGA---CCCTTGCCTGTCTGTCCAGGAATG 855  
Db 107 TCTCCATCTCTCATCAACCTGAACAGGAGACCGA---CCCTTGCCTGTCTGTCCAGGAATG 48  
Qy 856 TCAGTGTCTGTG 867  
Db 47 TCAGTGTCTGTG 36

RESULT 13  
AAS92200  
ID AAS92200 standard; cDNA; 463 BP.  
XX  
XX AAS92200;  
XX  
XX 13-FEB-2002 (first entry)  
DT  
XX  
XX DNA encoding novel human diagnostic protein #28004.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US0008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA

```
XX  Drmanac RT, Liu C, Tang YT;
PI  WPI: 2001-639362/73.
XX  P-PSDB; ABG28013.
XX  New isolated polynucleotide and encoded polypeptides, useful in
PT  diagnostics, forensics, gene mapping, identification of mutations
PT  responsible for genetic disorders or other traits and to assess
PT  biodiversity.
XX  Claim 1; SEQ ID NO 28004; 103pp; English.
XX  The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC  sequences. (I) is useful as hybridisation probes, polymerase chain
CC  reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC  and in recombinant production of (II). The polynucleotides are also used
CC  in diagnostics as expressed sequence tags for identifying expressed
CC  genes. (I) is useful in gene therapy techniques to restore normal
CC  activity of (II) or to treat disease states involving (II). (II) is
CC  useful for generating antibodies against it, detecting or quantitating a
CC  polypeptide in tissue, as molecular weight markers and as a food
CC  supplement. (II) and its binding partners are useful in medical imaging
CC  of sites expressing (II). (I) and (II) are useful for treating disorders
CC  involving aberrant protein expression or biological activity. The
CC  polypeptide and polynucleotide sequences have applications in
CC  diagnostics, forensics, gene mapping, identification of mutations
CC  responsible for genetic disorders or other traits to assess biodiversity
CC  and to produce other types of data and products dependent on DNA and
CC  amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC  coding sequences of the invention. Note: The sequence data for this
CC  patent did not appear in the printed specification, but was obtained in
CC  electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences
XX  Sequence 463 BP; 124 A; 133 C; 122 G; 84 T; 0 U; 0 Other;
XX  Query Match 10.3%; Score 195.6; DB 5; Length 463;
XX  Best Local Similarity 93.5%; Pred. No. 1.2e-42;
XX  Matches 215; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
XX  QY 1334 GGCTGGGCTCTTACAGCCAGAGATGCCCTCGGAAGCCTGGAGACGGACAGGA 1393
XX  Db 234 GTCTCAGCCCTGGCTGAACAAAGATGCCCTCGGAAGCCTGGAGACGGACAGGA 293
XX  QY 1394 CCAGAGGACCACTTCTCCCGATCCTGGACTTGGTACAGTGTGGAGCCAGCTCTCCA 1453
XX  Db 294 CCAGAGGACCACTTCTCCCGATCCTGGACTTGGTACAGTGTGGAGCCAGCTCTCCA 353
XX  QY 1454 GGCCACGGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCAGTGCCAGAGTCCATGACA 1512
XX  Db 354 GGCCACGGAAGTCTCTGAGCAGCCTGTTACTTCTGCTCAGTGCCAGAGTCCATGACA 413
XX  QY 1513 ATTAGTGAATGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 1562
XX  Db 414 ATTAGTGAATGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGA 463
XX  RESULT 14
XX  AAS73549/c
XX  ID AAS73549 standard; cDNA; 474 BP.
XX  AC AAS73549;
XX  XX 13-FEB-2002 (first entry)
XX  DT DNA encoding novel human diagnostic protein #9353.
XX  DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX  KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX  OS Homo sapiens.
XX  AC AAC11927;

FN WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG03362.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 9353; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX Sequence 474 BP; 75 A; 143 C; 122 G; 134 T; 0 U; 0 Other;
XX Query Match 7.1%; Score 135.4; DB 5; Length 474;
XX Best Local Similarity 79.6%; Pred. No. 3.2e-26;
XX Matches 160; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
XX QY 238 AGGTCACCTGTCTGTGTGTCGCCAGTTTGAAGCCGCTCTGCAGCATGCTTGAAGAGGAGT 297
XX Db 327 AGGTCACCTGTCTGTGTGTCGCCAGTTTGAAGCCGCTCTGCAGCATGCTTGAAGAGGAGT 268
XX QY 298 CGAGGATTGGCACTTCACAGCGGCGAGCCATCAAGCAGCAGCGGCTTTGCCAGCAAAACC 357
XX Db 267 CGAGGATTGGCACTTCACAGCGGCGAGCCATCAAGCAGCAGCGGCTTTGCCAGCAAAACC 208
XX QY 358 GAAACAGAGCCCGTGTCTTGTGTTACTACGTGAAGAGAGTCTCTCAACAAGCAGCAGCTGCAG 417
XX Db 207 GAAACAGGTGAACCTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGCTGGAG 148
XX QY 418 CGCTTCTACTCCCTGCGCCAC 438
XX Db 147 GAGGAGAACAGCCTGGGTGAC 127
XX RESULT 15
XX AAC11927
XX ID AAC11927 standard; cDNA; 190 BP.
XX XX
XX AC AAC11927;
```

XX 06-OCT-2000 (first entry)  
XX DT  
XX DE Human secreted protein 5' EST, SEQ ID NO: 16002.  
XX  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX gene therapy; chromosome mapping; ss.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-00200610.  
XX PR 26-FEB-1999; 99US-0122487P.  
XX PA (GEST ) GENSET.  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX PS Claim 1; SEQ ID NO 16002; 71pp + Sequence Listing; English.  
XX  
XX CC The present sequence is one of a large number of 5' ESTs derived from  
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
XX ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
XX to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors  
XX  
SQ Sequence 190 BP; 47 A; 42 C; 73 G; 28 T; 0 U; 0 Other;  
  
Query Match 7.0%; Score 133.4; DB 3; Length 190;  
Best Local Similarity 99.3%; Pred. No. 7.2e-26;  
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1578 CAGATCACTCGAAACCTGCTGCACGCTGAGATGGAGCACTCAGCGCGCTCCGCAAGA 1637  
Db |||||||  
19 CAGATCACTCGAAACCTGCTGCACGCTGAGATGGAGCACTCAGCGCGCTCCGCAAGA 78  
  
QY 1638 GGTGGACACCTTGAAGAAGAGGTGGCTGAACAGGAGGAGCGGCGCATGAAGTCCA 1697  
Db |||||||  
79 GGTGGACACCTTGAAGAAGAGGTGGCTGAACAGGAGGAGCGGCGCATGAAGTCCA 138  
  
QY 1698 GCGCTGGCCAGCTA 1712  
Db |||||||  
139 GCGCTGGCCAGGTA 153

Search completed: June 10, 2005, 06:10:07  
Job time : 655.812 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 724.339 Seconds  
(without alignments)  
16311.995 Million cell updates/sec

Title: US-09-155-676B-1

Perfect score: 1906

Sequence: 1 cattgggtcacggtggtggc.....tcattccttgataaaaaaaa 1906

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	263.8	13.8	280	9	US-09-563-817-201
2	215.8	11.3	542	16	US-10-029-386-8542
3	212.8	11.2	284	16	US-10-029-386-2242
4	128	6.7	500	16	US-10-029-386-8336
5	124	6.5	124	16	US-10-029-386-22036
6	63.8	3.3	3931	18	US-10-276-774-948
7	59.6	3.1	263	9	US-09-867-701-9696

8	54.6	2.9	371	10	US-09-918-995-24284	Sequence 24284, A
9	54.6	2.9	446	9	US-09-880-107-1518	Sequence 1518, Ap
10	54	2.8	464	10	US-09-918-995-15177	Sequence 15177, A
11	53.2	2.8	476	21	US-09-854-867-118	Sequence 118, App
12	53.2	2.8	476	21	US-10-786-970A-118	Sequence 118, App
13	50.4	2.6	7093	15	US-10-101-510-527	Sequence 527, App
14	50.4	2.6	10126	17	US-10-242-355-1024	Sequence 1024, Ap
15	50.2	2.6	550	19	US-10-437-963-63143	Sequence 63143, A
16	49.6	2.6	7140	21	US-10-211-028-3	Sequence 3, Appl
17	49.6	2.6	90597	21	US-10-211-028-1	Sequence 1, Appl
18	49.2	2.6	18535	9	US-09-764-878-385	Sequence 385, App
19	49.2	2.6	18535	14	US-10-079-854-385	Sequence 385, App
20	48	2.5	2463	13	US-10-027-632-102953	Sequence 102953,
21	48	2.5	2463	17	US-10-027-632-102953	Sequence 102953,
22	46.6	2.4	1182	13	US-10-027-348-1	Sequence 1, Appl
23	46.6	2.4	1185	13	US-10-027-348-13	Sequence 13, Appl
24	45.8	2.4	441	15	US-10-106-698-2578	Sequence 2578, Ap
25	45.6	2.4	962	20	US-10-723-860-6507	Sequence 6507, Ap
26	45	2.4	2033	9	US-09-764-868-623	Sequence 623, App
27	44.8	2.4	2056	17	US-10-151-469-13	Sequence 13, Appl
28	44.8	2.4	2056	19	US-10-717-597-106	Sequence 106, App
29	44.8	2.4	2056	19	US-10-802-432-14	Sequence 14, Appl
30	44.8	2.4	2056	20	US-10-723-860-2201	Sequence 2201, Ap
31	44.4	2.3	505	13	US-10-027-632-276057	Sequence 276057,
32	44.4	2.3	505	17	US-10-027-632-276057	Sequence 276057,
33	44	2.3	520	19	US-10-437-963-24663	Sequence 24663, A
34	44	2.3	550	9	US-09-925-298-379	Sequence 379, App
35	44	2.3	550	14	US-10-102-806-379	Sequence 379, App
36	44	2.3	4216	19	US-10-437-963-53652	Sequence 53652, A
37	43.8	2.3	1599	19	US-10-437-963-2771	Sequence 2771, Ap
38	43.6	2.3	531	9	US-09-925-298-381	Sequence 381, App
39	43.6	2.3	531	14	US-10-102-806-381	Sequence 381, App
40	43.4	2.3	1162	14	US-10-198-846-8983	Sequence 8983, Ap
41	43.2	2.3	505	13	US-10-027-632-276059	Sequence 276059,
42	43.2	2.3	505	17	US-10-027-632-276059	Sequence 276059,
43	43	2.3	850	18	US-10-425-114-29073	Sequence 29073, A
44	43	2.3	862	20	US-10-425-115-10220	Sequence 10220, A
45	43	2.3	940	9	US-09-789-561-23	Sequence 23, Appl

#### ALIGNMENTS

#### RESULT 1

US-09-563-817-201  
; Sequence 201, Application US/09563817  
; Patent No. US20020095031A1  
; GENERAL INFORMATION:  
; APPLICANT: Nehls, Michael C.  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: Polypeptides Encoded Thereby  
; FILE REFERENCE: LEX-0021-USA  
; CURRENT APPLICATION NUMBER: US/09/563,817  
; PRIOR FILING DATE: 2001-06-11  
; PRIOR APPLICATION NUMBER: US 60/132,343  
; PRIOR FILING DATE: 1999-05-04  
; NUMBER OF SEQ ID NOS: 1008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 201  
; LENGTH: 280  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(280)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-563-817-201

Query Match 13.8%; Score 263.8; DB 9; Length 280;  
Best Local Similarity 98.6%; Pred. No. 7.9e-71;  
Matches 276; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1464 GTCTCTGAGCAGCTGTTA-CTTCTGCTCAGTGCAGAGTCCATGACAAATTAGTGAAC 1522  
Db 1 GGCTATAGAGCAGCTGNHACCTTCTGCTCAGTGCAGAGTCCATGACAAATTAGTGAAC 60  
Qy 1523 TGGCCAGGCGCAGTGTGGCCATGATGAACAGGAAGATGAGCTGGAGGAGGAGAACAGAT 1582  
Db 61 TGGCCAGGCGCAGTGTGGCCATGATGAACAGGAAGATGAGCTGGAGGAGGAGAACAGAT 120  
Qy 1583 CACTGCGAACTCTCGACGGTGAGATGGAGCACTCAGCGCGCTCCGGCAAGAGGTGG 1642  
Db 121 CACTGCGAACTCTCGACGGTGAGATGGAGCACTCAGCGCGCTCCGGCAAGAGGTGG 180  
Qy 1643 ACACCTTGAAGAAAGTGGCTGAACAGGAGGAGCGGCGAGGCGATGAAGTCCAGGCGC 1702  
Db 181 ACACCTTGAAGAAAGTGGCTGAACAGGAGGAGCGGCGAGGCGATGAAGTCCAGGCGC 240  
Qy 1703 TGGCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 1742  
Db 241 TGGCAGCTATCTTTGCTATTTTGTGAGGAGATTCTAACC 280

RESULT 2  
US-10-029-386-8542/c  
; Sequence 8542, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 8542  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: NT HIT: g114779172, EVALUE 0.00e-00  
; OTHER INFORMATION: SWISSPROT HIT: 014197, EVALUE 5.90e-01  
; OTHER INFORMATION: EST\_HUMAN HIT: BG340557.1, EVALUE 0.00e+00  
US-10-029-386-8542

Query Match 11.3%; Score 215.8; DB 16; Length 542;  
Best Local Similarity 95.7%; Pred. No. 8.5e-56;  
Matches 244; Conservative 0; Mismatches 7; Indels 4; Gaps 2;  
Qy 614 CAGGTCTGAACTCCATCTCTTTGGGATTAACATGACAAAGGATTTGAACGGGCAGA 673  
Db 495 CAGGTCCGAACCTCCATCTCTTTGGGATTAACATGACAAAGGATTTGAACGGGCAGA 436  
Qy 674 GTAAGTTGCTCCACCGTTTCAGACCTCTTAAGGAGTCAACGCAAGAGTCAACCTCTCT 733  
Db 435 GTAAGTTGCTCCACCGTTTCAGACCTCTTAAGGAGTCAACGCAAGATGTGA---CCT 379  
Qy 734 TGCTGAAGGAGTCCACCAAGGAGTGAAGAGCTGTTCAGGAGATCAAGCTCTCTCTG 793  
Db 378 TGCTGAAGGAGTCCACCAAGGAGTGAAGAGCTGTTCAGGAGATCAAGCTCTCTCTG 319  
Qy 794 CCGTCTCCATCTCATCAACCTGAACAGGAGCGCA-CCCTTGCTGTGCTGCTCCAGGA 852  
Db 318 CCATCTCCATCTCATCAACCTGAACAGGAGCGCAACCCCTTGCTGCTGCTGCTCCAGGA 259  
Qy 853 ATGTCAGTGTGATG 867  
Db 258 ATGTCAGTGTGATG 244

RESULT 3  
US-10-029-386-22242/c  
; Sequence 22242, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 22242  
; LENGTH: 284  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: SWISSPROT HIT: 014197, EVALUE 2.10e-01  
; OTHER INFORMATION: NT HIT: g114779172, EVALUE 0.00e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: BF975008.1, EVALUE 0.00e+00  
US-10-029-386-22242

Query Match 11.2%; Score 212.8; DB 16; Length 284;  
Best Local Similarity 95.6%; Pred. No. 5.3e-55;  
Matches 241; Conservative 0; Mismatches 7; Indels 4; Gaps 2;  
Qy 617 GTCTGAACTCCATCTCTTTGGGATTAACATGACAAAGGATTTGAACGGGCAGATA 676  
Db 284 GTCTGAACTCCATCTCTTTGGGATTAACATGACAAAGGATTTGAACGGGCAGATA 225  
Qy 677 AGTTTGTCTCCACCGTTTCAGACCTCTTAAGGAGTCAACGCAAGAGTCAACCTCTCTTC 736  
Db 224 AGTTTGTCTCCACCGTTTCAGACCTCTTAAGGAGTCAACGCAAGATGTGA---CCTTC 168  
Qy 737 TGAAGGAGTCCACCAAGGAGTGAAGAGCTGTTCAGGAGATCAAGCTCTCTCTGCGG 796  
Db 167 TGAAGGAGTCCACCAAGGAGTGAAGAGCTGTTCAGGAGATCAAGCTCTCTCTGCGG 108  
Qy 797 TCTCCATCTCATCAACCTGAACAGGAGCGCA-CCCTTGCTGTGCTGCTCCAGGAATG 855  
Db 107 TCTCCATCTCATCAACCTGAACAGGAGCGCAACCCCTTGCTGCTGCTGCTCCAGGAATG 48  
Qy 856 TCAGTGTGATG 867  
Db 47 TCAGTGTGATG 36

RESULT 4  
US-10-029-386-8336  
; Sequence 8336, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 8336  
; LENGTH: 500  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
; OTHER INFORMATION: SWISSPROT HIT: Q92883, EVALUE 1.20e+00  
; OTHER INFORMATION: NT HIT: g115315787, EVALUE 2.00e-65  
; OTHER INFORMATION: EST\_HUMAN HIT: AL135642.1, EVALUE 2.00e-65  
US-10-029-386-8336

Query Match 6.7%; Score 128; DB 16; Length 500;  
Best Local Similarity 100.0%; Pred. No. 1.4e-28;  
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 237 CAGGGTCACCTGCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGGAG 296  
Db 341 CAGGGTCACCTGCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGGAG 400  
Qy 297 TCAGAGATTGGCACTACAGCGGCGGATCAAGCAGCGGCGCTTCCAGCAAAAC 356  
Db 401 TCAGAGATTGGCACTACAGCGGCGGATCAAGCAGCGGCGCTTCCAGCAAAAC 460  
Qy 357 CGAAACAG 364  
Db 461 CGAAACAG 468

RESULT 5  
US-10-029-386-22036  
; Sequence 22036, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 22036  
; LENGTH: 124  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR16.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
; OTHER INFORMATION: NT HIT: g115315787, EVALUE 9.00e-64  
; OTHER INFORMATION: EST\_HUMAN HIT: AL135642.1, EVALUE 1.00e-63  
; OTHER INFORMATION: SWISSPROT HIT: Q9SXU1, EVALUE 6.00e+00  
US-10-029-386-22036

Query Match 6.5%; Score 124; DB 16; Length 124;  
Best Local Similarity 100.0%; Pred. No. 1.3e-27;  
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 241 GTACCTGTCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGGAGTCA 300  
Db 1 GTACCTGTCTGTGTGCCAGTTTGAAGCCGCTCTGCAGCATGGCTTGAAGAGGAGTCA 60  
Qy 301 GGATTGGCACTACAGCGGCGGATCAAGCAGCGGCGCTTTCAGCAAAACCGAA 360  
Db 61 GGATTGGCACTACAGCGGCGGATCAAGCAGCGGCGCTTTCAGCAAAACCGAA 120  
Qy 361 ACAG 364  
|||||

Db 121 ACAG 124  
RESULT 6  
US-10-276-774-948/c  
; Sequence 948, Application US/10276774  
; Publication No. US20040053245A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang, Y. Tom et al  
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-030  
; CURRENT APPLICATION NUMBER: US/10/276,774  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 2700  
; SOFTWARE: Custom  
; SEQ ID NO 948  
; LENGTH: 3931  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-276-774-948

Query Match 3.3%; Score 63.8; DB 18; Length 3931;  
Best Local Similarity 51.6%; Pred. No. 3.3e-08;  
Matches 146; Conservative 0; Mismatches 137; Indels 0; Gaps 0;  
Qy 368 CCGTGTCTGGTACTACGTGAAGAGGTCCTCAACAGCAGCAGCTGCAGCGCTTCTACT 427  
Db 3721 CTGGCTACTGGGTGCTCGTGGTGCATTTTACTCGGAGAGGCGCATCAAGCAGATCGAGG 3662  
Qy 428 CCCTGCGCCACATCGCCTCAGAGCTGGCGGGGTCCGCGCTGGCTGGCTGCGCTCA 487  
Db 3661 TCGTCAGCAGCTGGCGCCACCACTTGGGCGCAGCGCTGCTGGCTGTACTCGGCCCTCA 3602  
Qy 488 ACGAACACTCCCTGGAGCGCTACTGCATCATGTCTCTGGCCGACCGCTGCAGCTGAGCA 547  
Db 3601 ACGAGAACTCCTTGGAGAGTACTCTCGGTGTGTTCCAGAGAACTTGGGCGCTGTCATA 3542  
Qy 548 CTTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCAGTATGCTTCTTACCA 607  
Db 3541 AGTACTACGTCAAGATGCCCTGGTCTGCAGCCACGATCACCTGACGCTCTTCTGACCT 3482  
Qy 608 TGCAGCAGCTCTGAATCCATCTCTTTGGCAGTTAACTCGA 650  
Db 3481 TGGTGTCCGGGCTAGAGTTCAITTCGTTTCGAGCTGGATCTGGA 3439

RESULT 7  
US-09-867-701-9696/c  
; Sequence 9696, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9696  
; LENGTH: 263  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-9696  
Query Match 3.1%; Score 59.6; DB 9; Length 263;

Query Match 2.9%; Score 54.6; DB 9; Length 446;  
Best Local Similarity 73.5%; Pred. No. 8.2e-06;  
Matches 83; Conservative 0; Mismatches 29; Indels 1; Gaps 1;  
QY 1795 TCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCTGTGTTTCCTCT 1852  
Db 136 GTCTTGTTGACATCTTTTGAGCCCTTAGACCAAGCTTTACCTGAAGCAGAGCTACCTCA 77  
QY 1853 GGACTTTTCATGTATGTGAGCCAATAAATGCTTTTCATCTTCAATGCTTGAAGCTGTGTTTCCTCT 1906  
Db 76 GAACCTTTTCAGTATGTGAGCCAATAAATGCTTTTCATCTTCAATGCTTGAAGCTGTGTTTCCTCT 23

RESULT 8  
US-09-918-995-24284  
; Sequence 24284, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 24284  
; LENGTH: 371  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-995-24284

Query Match 2.9%; Score 54.6; DB 10; Length 371;  
Best Local Similarity 73.5%; Pred. No. 7.4e-06;  
Matches 83; Conservative 0; Mismatches 29; Indels 1; Gaps 1;  
QY 1795 TCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCCCTGAAGCTGTGTTTCCTCTCTG 1853  
Db 247 TCTGATGACAGCGTTTGTGCCCCCTGGATCCAGCTGCTGAAGCTAGATATCCCTCTG 306  
QY 1854 GACTTTTCATGTATGTGAGCCAATAAATGCTTTTCATCTTCAATGCTTGAAGCTGTGTTTCCTCT 1906  
Db 307 GACTTTTCATGTATGTGAGCCAATAAATGCTTTTCATCTTCAATGCTTGAAGCTGTGTTTCCTCT 359

RESULT 9  
US-09-880-107-1518/c  
; Sequence 1518, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; APPLICANT: Gene Logic, Inc.  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1518  
; LENGTH: 446  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA621242  
US-09-880-107-1518

Query Match 2.9%; Score 54.6; DB 10; Length 464;  
Best Local Similarity 57.8%; Pred. No. 1.3e-05;  
Matches 96; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 365 AGCCCGTGTCTGGTACTAGTGAAGAGGTCTCTCAACAGCAGCAGCTGCAGGGTCTCT 424  
Db 298 AGCCTGTCTTGTGCCCCCTCTGAAAGCTGTCAACCCCAACACATCATCTCAGAGTTGG 357  
QY 425 ACTCCCTGCCACATCGCTCAGAGCTGGCGGGTCCGCTGCTGCTGCTGCTGCTGCTG 484  
Db 358 AGCACCTGAGCTTTGTCAACATGATGTGGCGGCTGCCGGGATGGCTGGCGTGGCCCC 417  
QY 485 TCAACGAACACTCCCTGGAGCGCTTACCTGCACATGCTTCCTGGCGGA 530  
Db 418 TGAACAATGGCGCTGATGGAGTGTCTACCTGAAGCTGTCTGCTGCAGGA 463

RESULT 10  
US-09-918-995-15177  
; Sequence 15177, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 15177  
; LENGTH: 464  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-995-15177

Query Match 2.8%; Score 54; DB 10; Length 464;  
Best Local Similarity 57.8%; Pred. No. 1.3e-05;  
Matches 96; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 365 AGCCCGTGTCTGGTACTAGTGAAGAGGTCTCTCAACAGCAGCAGCTGCAGGGTCTCT 424  
Db 298 AGCCTGTCTTGTGCCCCCTCTGAAAGCTGTCAACCCCAACACATCATCTCAGAGTTGG 357  
QY 425 ACTCCCTGCCACATCGCTCAGAGCTGGCGGGTCCGCTGCTGCTGCTGCTGCTGCTG 484  
Db 358 AGCACCTGAGCTTTGTCAACATGATGTGGCGGCTGCCGGGATGGCTGGCGTGGCCCC 417  
QY 485 TCAACGAACACTCCCTGGAGCGCTTACCTGCACATGCTTCCTGGCGGA 530  
Db 418 TGAACAATGGCGCTGATGGAGTGTCTACCTGAAGCTGTCTGCTGCAGGA 463

RESULT 11  
US-09-854-867-118  
; Sequence 118, Application US/09854867  
; Publication No. US20030224356A1  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL H  
; APPLICANT: ROGAN, PETER K  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/09/854,867  
; CURRENT FILING DATE: 2003-05-08  
; NUMBER OF SEQ ID NOS: 613  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 118  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat\_region  
; LOCATION: (1)..(476)  
; OTHER INFORMATION:  
US-09-854-867-118



Query Match 2.8%; Score 53.2; DB 10; Length 476;  
Best Local Similarity 69.5%; Pred. No. 2.3e-05;  
Matches 89; Conservative 0; Mismatches 33; Indels 6; Gaps 1;

QY 1776 GAGAGAGAAATCCACAGTTCTCGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCT 1835  
DB 313 GAGATGAGAGAGACAGATTCTCGATGACATCATTTTGAGCACCTGGATCCAGCCATGCCT 372  
QY 1836 GAAG-----CTGTGTTTCTCTGGACTTTTCATGTATGTAGAGCCCAATAAATGCTTTCA 1889  
DB 373 GAAGCCAGATCTATGCTACCCCTGGACTTTTCAGTTACATGAGCCCAATAAATCTCTTTT 432  
QY 1890 TTCCTTGA 1897  
DB 433 TTGCTTAA 440

RESULT 12  
US-10-786-970A-118  
; Sequence 118, Application US/10786970A  
; Publication No. US200506449A1  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATING  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/10/786, 970A  
; CURRENT FILING DATE: 2004-02-24  
; PRIOR APPLICATION NUMBER: US/09/573, 080  
; PRIOR FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 118  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat region  
; LOCATION: (1)..(476)  
; PUBLICATION INFORMATION:  
; PUBLICATION INFORMATION:  
; AUTHORS: Jurka, J; Wallchiewicz, J; Milosavljevic, A  
; TITLE: Prototypic sequences for human repetitive DNA  
; JOURNAL: Journal of Molecular Evolution  
; VOLUME: 35  
; ISSUE: 4  
; PAGES: 286-291  
; DATE: 1992-10-  
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
; DATABASE ENTRY DATE:  
; DATABASE ENTRY DATE: 1996-01-26  
US-10-786-970A-118

Query Match 2.8%; Score 53.2; DB 21; Length 476;  
Best Local Similarity 69.5%; Pred. No. 2.3e-05;  
Matches 89; Conservative 0; Mismatches 33; Indels 6; Gaps 1;

QY 1776 GAGAGAGAAATCCACAGTTCTCGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCT 1835  
DB 313 GAGATGAGAGAGACAGATTCTCGATGACATCATTTTGAGCACCTGGATCCAGCCATGCCT 372  
QY 1836 GAAG-----CTGTGTTTCTCTGGACTTTTCATGTATGTAGAGCCCAATAAATGCTTTCA 1889  
DB 373 GAAGCCAGATCTATGCTACCCCTGGACTTTTCAGTTACATGAGCCCAATAAATCTCTTTT 432  
QY 1890 TTCCTTGA 1897  
DB 433 TTGCTTAA 440

RESULT 13  
US-10-101-510-527

; Sequence 527, Application US/10101510  
; Publication No. US20030148295A1  
; GENERAL INFORMATION:  
; APPLICANT: WAN, JACKSON  
; APPLICANT: WANG, YIXIN  
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
; FILE REFERENCE: 15117.0012  
; CURRENT APPLICATION NUMBER: US/10/101,510  
; CURRENT FILING DATE: 2002-03-20  
; PRIOR APPLICATION NUMBER: 60/276,947  
; PRIOR FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 805  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 527  
; LENGTH: 7093  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-101-510-527

Query Match 2.6%; Score 50.4; DB 15; Length 7093;  
Best Local Similarity 74.0%; Pred. No. 0.00065;  
Matches 77; Conservative 0; Mismatches 26; Indels 1; Gaps 1;  
QY 1795 TCCTGATAGTCTCATTTGAGCTCCTGGATCCAGTCTTTCTGAGCT-GTGTTTCTCTG 1853  
DB 6443 TCCTGATGACAGCGTTTGTGCCCTCGATCCAAACCGTGCCTGAAGCTAGAAATATCCCTG 6502  
QY 1854 GACTTTTCATGTATGTGAGCCCAATAAATGCTTTTCAATCTCTTGA 1897  
DB 6503 GACTTTTCAGTTATGTGACCAATAAATACCCCTTTTTCGCTTAA 6546

RESULT 14  
US-10-242-355-1024  
; Sequence 1024, Application US/10242355  
; Publication No. US20030235831A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC003C1  
; CURRENT APPLICATION NUMBER: US/10/242,355  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 09/764,897  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1267  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1024  
; LENGTH: 10126  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-242-355-1024  
Query Match 2.6%; Score 50.4; DB 17; Length 10126;  
Best Local Similarity 74.0%; Pred. No. 0.00077;

Mon Jun 13 13:27:09 2005

Matches 77; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 1795 TCCTGATAGTCTCATTTGAGCTCTGATCCAGTCTTTCTGAGCT-GTGTTCCTCTG 1853  
|||||  
Db 9473 TCCTGATGACAGCGTTTGTGCCCTGGATCAACCGTGCCTGAGCTAGAAATATCCCTG 9532  
|||||

QY 1854 GACTTTTCATGTATGTGAGCCATAAATGCTTTCATTCCTTGA 1897  
|||||  
Db 9533 GACTTTTCAGTTAIGTGAACCAATAATACCTTTTGTCTAA 9576  
|||||

RESULT 15

US-10-437-963-63143

; Sequence 63143, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 63143

; LENGTH: 550

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(550)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_64413C.1

US-10-437-963-63143

Query Match 2.6%; Score 50.2; DB 19; Length 550;

Best Local Similarity 87.3%; Pred. No. 0.00021;

Matches 55; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 9 CACGCGGTGGCGGCTCTAGATAGTGGATCCCGGGCTGCAGGAATTCGATTCGAGG 68  
|||||

Db 21 CACCGGTGGCGGCTCTAGATAGTGGATCCCGGGCTGCAGGAATTCGACGACA 80  
|||||

QY 69 CCA 71  
|||

Db 81 GCA 83

Search completed: June 11, 2005, 06:12:51

Job time : 726.339 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 4016.07 Seconds  
(without alignments)  
18065.048 Million cell updates/sec

Title: US-09-155-676B-1  
Perfect score: 1906  
Sequence: 1 cattgggtacacgggtggcg.....tcattcttggaaaaaaa 1906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	797	41.8	1981	3	BC024293
2	672.2	35.3	730	1	AU122438
3	648	34.0	691	4	BM668241
4	629.8	33.0	998	5	BQ064251
5	627.6	32.9	990	5	BQ060776
6	624.2	32.7	970	5	BQ065026
7	610.2	32.0	829	4	BG757448
8	603.8	31.7	835	4	BG758677
9	603.8	31.7	1022	5	BQ057616
10	600.6	31.5	843	4	BG758425
11	600.4	31.5	892	5	BQ037008
12	586	30.7	609	4	BM692924
13	550.8	28.9	813	4	BG684262
14	522.6	27.4	626	5	BQ184171
15	513.2	26.9	819	4	BG759681
16	509.6	26.7	620	5	BQ185857
17	489	25.7	769	4	BG684980
18	482.8	25.3	504	4	BM697878
19	482	25.3	914	4	BG340557
20	481.6	25.3	782	4	BG759163
21	475.4	24.9	951	4	BG685372
22	474.4	24.9	744	9	AY399018
23	467.6	24.5	1064	5	BQ065934
24	461	24.2	501	1	AL135642

25	454.8	23.9	737	4	BF975008
26	454.2	23.8	1075	5	BQ058038
27	453	23.8	855	5	BX760749
28	447	23.5	1051	2	BQ061744
29	437.2	22.9	576	2	AW974284
30	436	22.9	965	4	BG757633
31	431.4	22.6	822	4	BM478075
32	427	22.4	439	1	AI369689
33	410.8	21.6	682	9	AY399019
34	385.4	20.2	744	9	AY399020
35	382.4	20.1	419	1	AA593354
36	376.8	19.8	850	5	BQ642454
37	372.8	19.6	430	2	AW043206
38	372.4	19.5	522	2	AW044949
39	370	19.4	506	5	BX283539
40	358.8	18.8	520	2	AW503340
41	345	18.1	847	7	CR579211
42	343	18.0	478	5	BX283759
43	341.6	17.9	646	1	AL965834
44	323.2	17.0	602	1	AL959249
45	317.6	16.7	498	4	BG759800

#### ALIGNMENTS

RESULT 1  
BC024293  
LOCUS BC024293 1981 bp mRNA linear HTC 04-MAR-2003  
DEFINITION Homo sapiens, clone IMAGE:5088007, mRNA.  
ACCESSION BC024293  
VERSION BC024293.1 GI:22028130  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1981)  
AUTHORS Strausberg,R.  
TITLE Direct Submission  
JOURNAL Submitted (26-PEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada  
info@bcgsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 42 Row: m Column: 22  
This clone has the following problem: retained intron.

FEATURES  
source  
1. .1981  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5088007"  
/tissue\_type="Lymph, lymphoma"

/clone lib="NIH MGC_99"		AUI22438 MAMMA1 Homo sapiens cDNA clone MAMMA1002380 5', mRNA	
/lab host="DH10B"		sequence.	
/note="Vector: pOTB7"		AUI22438	
		AUI22438.1 GI:10937708	
		VERSION	
		EST.	
		SOURCE	
		Homo sapiens (human)	
		ORGANISM	
		REFERENCE	
		AUTHORS	
		TITLE	
		JOURNAL	
		COMMENT	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="MAMMA1002380"	
		/tissue_type="mammary gland"	
		/clone_lib="MAMMA1"	
		/note="Vector: pME18SFL3"	
		Location/Qualifiers	
		HRI human cDNA project	
		Unpublished (2000)	
		Contact: Takao Isogai	
		Genomics Laboratory	
		Helix Research Institute	
		1532-3 Yana, Kisarazu, Chiba 292-0812, Japan	
		Tel: 81-438-52-3975	
		Fax: 81-438-52-3986	
		Email: genomics@hri.co.jp	
		HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix	
		Research Institute; cDNA library construction: Department of	
		Virology, Institute of Medical Science, University of Tokyo, and	
		Helix Research Institute.	
		FEATURES	
		source	
		1. 730	
		/organism="Homo sapiens"	

QY 557 AAGACTGGTCTTTTGTGATGATGAAGAAAGGTCAGTATGCTTCTACCATGCGACGAG 616  
Db 502 AAGACTGGTCTTTTGTGATGATGAAGAAAGGTCAGTATGCTTCTACCATGCGACGAG 561  
QY 617 GTCTGAAGTCTCATACTCTTTTGGATTAACATCGACAAAGGATTTGAACGGGCGAGGTA 676  
Db 562 GTCTGAAGTCTCATACTCTTTGCGATTAACTCGACACACAGGATTTGACGGGCGAGGTA 621  
QY 677 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGACGAGAACTGACCTCTTTC 736  
Db 622 AGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGACGAGAACTGACCTCTTTC 680  
QY 737 TGAAGGAGTCCACCAAGGAGTGAAGCCTGTTTTCAGGAGATCACAGCCT 787  
Db 681 TGAAGGAGTCCACCAAGGAGTGANCA-CCTGTTTCAAGGAGATCACAGCCT 730

## RESULT 3

BM668241/c

LOCUS

DEFINITION

UI-E-CK1-afk-h-18-0-UI.s2 UI-E-CK1 Homo sapiens cDNA clone

UI-E-CK1-afk-h-18-0-UI 3', mRNA sequence.

BM668241

BM668241.1 GI:18976072

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 691)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1..691

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-CK1-afk-h-18-0-UI"

/tissue\_type="Retina Foveal and Macular"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-CK1"

/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-CK1 is a normalized cDNA library containing the

following tissue(s): Retina Foveal and Macular. The

library was constructed according to Bonaldo, Lennon and

Soares, Genome Research, 6:791-806, 1996. First strand

cDNA synthesis was primed with an oligo-dT primer

containing a Not I site. Double stranded cDNA was ligated

to an EcoR I adaptor, digested with Not I, and cloned

directionally into pT7T3-Pac vector. The oligonucleotide

used to prime the synthesis of first-strand cDNA contains

a library tag sequence that is located between the Not I

and the (dT)18 tail. The sequence tag for this

library is GTCC. This library was created for the program,

Gene Discovery in the Visual System, supported by National  
Eye Institute (NEI).  
TAG TISSUE=Foveal and Macular Retina  
TAG\_LIB=UI-E-CK1  
TAG\_SEQ=GTCC"

## ORIGIN

Query Match 34.0%; Score 648; DB 4; Length 691;  
Best Local Similarity 99.6%; Pred. No. 3.1e-164;  
Matches 681; Conservative 0; Mismatches 0; Indels 3; Gaps 3;  
QY 1225 GGAGTGGCCCGAGAGCCACTGGAAGGAAACACCTGCTCTCCAGATGCACAGTGGGC 1284  
Db 691 GGAGTGGCCCGAGAGCCACTGGAAGGAAACACCTGCTCTCCAGATGCACAGTGGGC 632  
QY 1285 TCCGCTGMAAGTCTGTCACAAATGACTCCGACATCTCTTCCCTGTCAGTGGCGTGGCTC 1344  
Db 631 TCCGCTGMAAGTCTGTCACAAATGACTCCGACATCTCTTCCCTGTCAGTGGCGTGGCTC 572  
QY 1345 CTACAGCCCGAGAGTGGCCCTCGGAAGCCTGGAGAACGGGACAGACGAGAGACCA 1404  
Db 571 CTACAGCCCGAGAGTGGCCCTCGGAAGCCTGGAGAACGGGACAGACGAGAGACCA 512  
QY 1405 CGTTCTCCCGAGTCTCGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCAGGCCACGGAAG 1464  
Db 511 CGTTCTCCCGAGTCTCGGACTTCGGTACAGTGTGGAAGCCAGCTCTCCAGGCCACGGAAG 452  
QY 1465 TCCTCTGAGCAGCCTGTGA-CCTTCTGCTCAGTCCGAGAGTCCATGACAAATTAGTGAAT 1523  
Db 451 TCCTCTGAGCAGCCTGTGTACTCTTCTGCTCAGTCCGAGAGTCCATGACAAATTAGTGAAT 392  
QY 1524 GCGCCAGGCCACTGTGGCCATGATGAACAGAGGAGTGTGGAGGAGGAGAACAGATC 1583  
Db 391 GCGCCAGGCCACTGTGGCCATGATGAACAGAGGAGTGTGGAGGAGGAGAACAGATC 332  
QY 1584 ACTGCGAAACCTGCTCGACGGTGAGATGGAGACTCAGCCGCTCCGCGAAGAGGTGGA 1643  
Db 331 ACTGCGAAACCTGCTCGACGGTGAGATGGAGACTCAGCCGCTCCGCGAAGAGGTGGA 272  
QY 1644 CACCTTGAAGAGGAGTGGCTGAAACAGAGGAGCGGAGGATGAGTCCAGGCGCT 1703  
Db 271 CACCTTGAAGAGGAGTGGCTGAAACAGAGGAGCGGAGGATGAGTCCAGGCGCT 212  
QY 1704 GGCAGCTATCTTTGCTATTTTGTGAGGAGATTTAAACCCACGCTGAGAACCATGTGCTG 1763  
Db 211 GGCAGCTATCTTTGCTATTTTGTGAGGAGATTTAAACCCACGCTGAGAACCATGTGCTG 152  
QY 1764 GAGAAATGGAGGAGAGAGAAATCCAAACAGTTCCTGTATGCTCTCATTGTAGCTCCTGGAT 1823  
Db 151 GAGAAATGGAGGAGAGAGAAAT-CAACAGTTCCTGTATGCTCTCATTGTAGCTCCTGGAT 93  
QY 1824 CAGTCTTTCTGAGAGCTGTGTTTCTCTGAGACTTTTTCATGTATGTGAG-CCATAAATT 1882  
Db 92 CAGTCTTTCTGAGAGCTGTGTTTCTCTGAGACTTTTTCATGTATGTGAGCCCAATAAATT 33  
QY 1883 GCTTTTCATTCTTTGAAAAAATAA 1906  
Db 32 GCTTTTCATTCTTTGAAAAAATAA 9

## RESULT 4

BM664251

LOCUS

DEFINITION

AGENCOURT\_6854182 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5925960

5', mRNA sequence.

BM664251

BM664251.1 GI:19892719

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 998)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LNCM2098 row: i column: 01  
High quality sequence stop: 692.  
Location/Qualifiers

FEATURES  
source

1..998  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5925960"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH-MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH-MGC  
Library."

ORIGIN

Query Match 33.0%; Score 629.8; DB 5; Length 998;  
Best Local Similarity 96.7%; Pred. No. 3e-159;  
Matches 675; Conservative 0; Mismatches 18; Indels 5; Gaps 3;

QY 544 AGCAGCTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAGCTCCAGTATGCTTCCT 603  
DB 92 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAGCTCCAGTATGCTTCCT 151  
QY 604 ACATGGCAGAGGTCTGAATCCATACCTCTTTTGGATTAACATCGACACAGAGATTG 663  
DB 152 ACCATGGCAGAGGTCCGAATCCATACCTCTTTTGGATTAACATCGACACAGAGATTG 211  
QY 664 AAGGGCAGAGTAAGTTTCTCCACCGTTTCAGACTCTTAAAGAGTCAACGAGAAC 723  
DB 212 AAGGGCAGAGTAAGTTTCTCCACCGTTTCAGACTCTTAAAGAGTCAACGAGAAC 271  
QY 724 GTGACCTCTTGTGAAGAGTCCACGCAAGAGTGAAGAGCTGTTCAGGAGATCA 783  
DB 272 GTGA---CCCTGTGAAGAGTCCACGCAAGAGTGAAGAGTGTTCAGGAGATCA 328  
QY 784 GCCTCTCTGCGCTCTCCATCTCATCAACCTGAACGAGGACCGA-CCCTTGTCTGTC 842  
DB 329 GCCTCTCTGCGCTCTCCATCTCATCAACCTGAACGAGGACCGA-CCCTTGTCTGTC 388  
QY 843 GTGTCCAGGAATCTCAGTGTCTGATGCAATGCAAAAGAGGAGCGAAGAAAGAA 902  
DB 389 GTGTCCAGGAATCTCAGTGTCTGATGCAATGCAAAAGAGGAGCGAAGAAAGCA 448  
QY 903 GTGACCAACAATAATCTCATTTGATGATGAAGAGATGAGCAGAACTCTCGGAGCTGTTT 962  
DB 449 GTGACCAACAATAATCTCATTTGATGATGAAGAGATGAGCAGAACTCTCGGAGCA 508  
QY 963 AAAAGACACCTGGGCGAGGAGCTCAGAGCAAACTCCGACCGCTCTCTCTGTCAT 1022  
DB 509 AAAAGACACCTGGGCGAGGAGCTCAGAGCAAACTCCGACCGCTCTCTCTGTCAT 568  
QY 1023 ATCATGTCCGCTTTTGAAGACCCCTTCGGGCTTAACCTCAATGAA-TCAGAGCAAC 1081

Db 569 ATCAATGTCGCTTTGAAGACCCCTTCGGGCCAAATCCCAATGGAAGTCAGACGACAC 628  
QY 1082 TCATGAAATTAATGATTCCTCTTTTGAACGGGAGTTTGGTACAGAGCTTGTATG 1141  
Db 629 TCGTGAATTAATGATTCCTCTTTTGAACGGGAGTTTGGTACAGAGCTTGTATG 688  
QY 1142 AAAGCATGATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1201  
Db 689 AAAGCATGATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 748  
QY 1202 GGAAGCAAGCAGAGGCGCACTCGGAGTCCGCCGAGAA 1239  
Db 749 GGACGAGCAGACAGGCGCACTCAGAGTCCGCCGAGAA 786

RESULT 5  
BQ060776 990 bp mRNA linear EST 02-APR-2002  
LOCUS AGENCOURT 7049038 NIH\_MGC\_99 Homo sapiens cDNA clone IMAGE:5816452  
DEFINITION 5' mRNA sequence.  
ACCESSION BQ060776  
VERSION BQ060776.1 GI:19884191  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 998)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaaps-remail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LNCM2073 row: f column: 05  
High quality sequence stop: 686.  
Location/Qualifiers

FEATURES  
source

1..990  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH-MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH-MGC  
Library."

ORIGIN

Query Match 32.9%; Score 627.6; DB 5; Length 990;  
Best Local Similarity 96.6%; Pred. No. 1.2e-158;  
Matches 674; Conservative 0; Mismatches 19; Indels 5; Gaps 3;

QY 544 AGCAGCTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAGCTCCAGTATGCTTCCT 603  
DB 98 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAGCTCCAGTATGCTTCCT 157  
QY 604 ACCATGGCAGAGGTCTGAATCCATACCTCTTTTGGATTAACATCGACACAGAGATTG 663

```
Db 158 ACCATGGCAGCAGGTCGGAACCTCCATACCTCTTTGCGATTAACTTGACAACAGGATTTG 217
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGAGTCAACGCAAC 723
Db 218 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGAGTCAACGCAAT 277
Qy 724 GTGACCTCTCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGAGCTGTTTCAGGAGATCACA 783
Db 278 GTGA---CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGAGCTGTTTCAGGAGATCACA 334
Qy 784 GCCTCTCTGCGCTCTCCATCTCATCAAACTCAACAGGAGACCGA-CCTTTGCTCTGTC 842
Db 335 GCCTCTCTGCGCTCTCCATCTCATCAAACTCAACAGGAGACCGA-CCTTTGCTCTGTC 394
Qy 843 GTGTCCAGGAATGTCACTGTCTGATGCCAAATGCAAAAGAGCGGAAAGAAAGAAA 902
Db 395 GTGTCCAGGAATGTCACTGTCTGATGCCAAATGCAAAAGAGCGGAAAGAAAGAAA 454
Qy 903 GTGACCAACATATCTCATTTGATGAGGAGATGAGCAGAACTCTGGGAGCTGTTT 962
Db 455 GTGACCAACATATCTCATTTGATGAGGAGATGAGCAGAACTCTGGGAGCATGTTT 514
Qy 963 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 515 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCACTCTCTGTCAAT 574
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAAATGGAA-TCAGAGCAGCAAC 1081
Db 575 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAAATGGAA-TCAGAGCAGCAAC 634
Qy 1082 TCATGGAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACGAAAGCTTGATGTG 1141
Db 635 TCGTGGAAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACGAAAGCTTGATGTG 694
Qy 1142 AAAAGCATCATGATGAAGATGTGATGAAACGAAAGTCACTGATGTAAGTCAATCA 1201
Db 695 AAAAGCATCATGATGAAGATGTGATGAAACGAAAGTCACTGATGTAAGTCAATCA 754
Qy 1202 GGAAGGAAGCAGAGGGGCCACTCGGAGTCGCGCCAGAA 1239
Db 755 AGACGGAAGCAGAGGGGCCACTCAGAGTCGCGCCAGAA 792
```

```
RESULT 6
BQ065026
LOCUS
DEFINITION AGENCOURT_6856627 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5929505
5', mRNA sequence.
```

```
ACCESSION BQ065026
```

```
VERSION BQ065026.1 GI:19894072
```

```
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
```

```
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
1 (bases 1 to 970)
```

```
NIH-MGC http://mgi.nci.nih.gov/
```

```
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
```

```
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-x@mail.nih.gov
```

```
Tissue Procurement: Lou Staudt
```

```
cDNA Library Preparation: Rubin Laboratory
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
```

```
DNA sequencing by: Agencourt Bioscience Corporation
```

```
Clone distribution: MGC clone distribution information can be
```

```
found through the I.M.A.G.E. Consortium/LNL at:
```

```
http://image.lnl.gov
```

```
Plate: LLCM2107 row: 1 column: 18
```

```
High quality sequence stop: 676.
```

```
Location/Qualifiers
```

```
1..970
```

```
/organism="Homo sapiens"
```

```
FEATURES
source
```

```
RESULT 7
BG757448
LOCUS
```

```
BG757448
```

```
829 bp
```

```
mRNA
```

```
linear
```

```
EST 15-MAY-2001
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5929505"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph. Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
```

```
ORIGIN
```

```
Query Match 32.7%; Score 624.2; DB 5; Length 970;
```

```
Best Local Similarity 96.7%; Pred. No. 9.9e-158;
```

```
Matches 670; Conservative 0; Mismatches 18; Indels 5; Gaps 3;
```

```
Qy 544 AGCAGCTTTTATGAAGAGTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTCCT 603
```

```
Db 92 AGTACTTTTATGAAGAGTGGTCTTTTGTGATGATGAAGAGGTCCAGTATGCTTCCT 151
```

```
Qy 604 ACCATGGCAGCAGGTCTGAATCCATCTCTTTGCGATTAACTCGAACAAGGATTTG 663
```

```
Db 152 ACCATGGCAGCAGGTCCGAACCTCATCTCTTTGCGATTAACTTGACAACAAGGATTTG 211
```

```
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTAAAGGAGTCAACGCAAC 723
```

```
Db 212 AACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACTCTTAAAGGAGTCAACGCAAT 271
```

```
Qy 724 GTGACCTCTTGTCTGAAGAGTCCACGCAAGGAGTGAAGAGTCTTTCAGGAGATCACA 783
```

```
Db 272 GTGA----CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGAGTCTTTCAGGAGATCACA 328
```

```
Qy 784 GCCTCTCTGCGCTCTCCATCTCTCATCAAACTGAAACAGGAGACCGA-CCTTTGCTCTGTC 842
```

```
Db 329 GCCTCTCTGCGCTCTCCATCTCTCATCAAACTGAAACAGGAGACCGA-CCTTTGCTCTGTC 388
```

```
Qy 843 GTGTCCAGGAATGTCACTGTCTGATGCCAAATGCAAAAGAGCGGAAAGAAAGAAA 902
```

```
Db 389 GTGTCCAGGAATGTCACTGTCTGATGCCAAATGCAAAAGAGCGGAAAGAAAGAAA 448
```

```
Qy 903 GTGACCAACATATCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGAGCTGTTT 962
```

```
Db 449 GTGACCAACATATCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGAGCATGTTT 508
```

```
Qy 963 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
```

```
Db 509 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCACTCTCTGTCAAT 568
```

```
Qy 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAAATGGAA-TCAGAGCAGCAAC 1081
```

```
Db 569 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAAATGGAAATGAGAGCAGCAAC 628
```

```
Qy 1082 TCATGGAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACGAAAGCTTGATGTG 1141
```

```
Db 629 TCGTGGAAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACGAAAGCTTGATGTG 688
```

```
Qy 1142 AAAAGCATCATGATGAAGATGTGATGAAACGAAAGTCACTGATGTAAGTCAATCA 1201
```

```
Db 689 AAAAGCATCATGATGAAGATGTGATGAAACGAAAGTCACTGATGTAAGTCAATCA 748
```

```
Qy 1202 GGAAGGAAGCAGAGGGGCCACTCGGAGTCGCGCC 1234
```

```
Db 749 GGAAGGAAGCAGAGGGGCCACTCAAAGTCGCCC 781
```

DEFINITION 602711092F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4851502 5',  
mRNA sequence.  
ACCESSION BG757448  
VERSION BG757448.1 GI:14068101  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 829)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCMI894 row: 9 column: 23  
High quality sequence stop: 801.  
Location/Qualifiers  
1. 829  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4851502"  
/tissue\_type="Primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_48"  
/note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

FEATURES  
source  
Query Match 32.0%; Score 610.2; DB 4; Length 829;  
Best Local Similarity 95.9%; Pred. No. 6e-154;  
Matches 670; Conservative 0; Mismatches 23; Indels 6; Gaps 4;

QY 544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCCT 603  
DB 133 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGGTCCAGTATGCTTCCT 192

QY 604 ACCATGGCAGGAGTCTGAATCCATATCTTTTGGGATTAAATCATCGACAAACAAGATTG 663  
DB 193 ACCATGGCAGGAGTCCCAACTCCATATCTTTTGGGATTAAATCATCGACAAACAAGATTG 252

QY 664 AACGGCAGAGTAAAGTTTGGTCCACCGTTTCAGACCTTTAAAGGAGTCAACGAGAAC 723  
DB 253 AACGGCAGAGTAAAGTTTGGTCCACCGTTTCAGACCTTTAAAGGAGTCAACGAGAAAT 312

QY 724 GTGACCTCTCTGCTGAAGAGTCCACGAAAGAGTGAAGAGTCTTTTCAGGAGATCACA 783  
DB 313 GTGA---CTTGTGAAGAGTCCACGCAAGAGTGAAGAGTGTTCAGGAGATCACA 369

QY 784 GCCTCCTCTGCGTCTCATCTCTATCAAACTGAAACAGGAGACGA-CCCTTGCCTGTC 842  
DB 370 GCCTCCTCTGCGATCTCCATCTCATCAAACTGAAACAGGAGACGACCCCTTGCCTGTC 429

QY 843 GTGTCCAGGAATGTCAAGTCTGATGCCAAATGCAAAAGAGGCGGAAAGAAAGAA 902  
DB 430 GTGTCCAGGAATGTCAAGTCTGATGCCAAATGCAAAAGAGGCGGAAAGAAAGCAA 489

QY 903 GTGACCAACATAATCTCTATTTTGTATGATGAGGAAGATGACAGAACTCTGGGGACGTGTTT 962

DB 490 GTGACCAACATATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT 549

QY 963 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGACAACTCCGACCGTCTCTGTCAAT 1022

DB 550 AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGACAACTCCGACCACTCTCTGTCAAT 609

QY 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAATGGAA-TCAGAGCAGCAAC 1081

DB 610 ATCATGTCCGCTTTGAAAGCCCTTCGGGCCTAACTCCAATGGAAGTCAGAGCAGCAAC 669

QY 1082 TCATGGAAAATTGATTCCTCTGTTTGAACGGGAGTTCGGGTACCAAGACCTTGTATGTG 1141

DB 670 TCGTGGAAAATTGATTCCTCTGTTTGAACAGGAGTTCGGGTACCAAGACCTTGTATGTG 729

QY 1142 AAAAGCATCCGATGATGAAGATGCGATGAAACCAAGATGACGTATGG-AAACTCATC 1200

DB 730 AAAAGCATCCGATGATGAAGATGCGATGAAACCAAGATGACGTATGGAAAACACTCATC 789

QY 1201 AGGAAGGAAGCACAGGGGCCACTCGAGTCGCCCGAGAA 1239

DB 790 AGGACGGAAGCACATGTGGCACTCAGAGTCGGGCGAGAA 828

RESULT 8  
BG758677 835 bp mRNA linear EST 15-MAY-2001  
LOCUS 602712923F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4853188 5',  
mRNA sequence.  
ACCESSION BG758677  
VERSION BG758677.1 GI:14069330  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 835)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCMI698 row: n column: 05  
High quality sequence stop: 819.  
Location/Qualifiers  
1. 835  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4853188"  
/tissue\_type="Primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_48"  
/note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

FEATURES  
source  
Query Match 31.7%; Score 603.8; DB 4; Length 835;  
Best Local Similarity 96.3%; Pred. No. 3.3e-152;

ORIGIN



```
Matches 673; Conservative 0; Mismatches 17; Indels 9; Gaps 5;
QY 544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCAGATGCTTCCT 603
Db |||
31 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGAGGTCAGATGCTTCCT 90
QY 604 ACCATGGCAGCAGGTCTGAACCTCATCTCTTTGGCATTAACATCAGCAACAGGATTTG 663
Db |||
91 ACCATGGCAGCAGGTCCGAACCTCATCTCTTTGGCATTTTGAAGAGATCAACGAGAT 147
QY 664 AACGGGAGAGTAAAGTTTGTCTCCACGTTTTCAGACTCTTTAAAGAGTCAACGAGAAC 723
Db |||
148 AACGGGAGAGTAAAGTTTGTCTCCACGTTTTCAGACTCTTTAAAGAGTCAACGAGAT 207
QY 724 GTGACCTCTCTGTGAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA 783
Db |||
208 GTGA---CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACA 264
QY 784 GCCTCTCTGCGTCTCCATCTCTCAATCAAACTGAACAGAGAGACCGA-CCCTTGCTGTG 842
Db |||
265 GCCTCTCTGCGTCTCCATCTCTCAATCAAACTGAACAGAGAGACCGACCCCTTGCCCGTC 324
QY 843 GTCTCAGGAATCTCAGTGTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGAA 902
Db |||
325 GTCTCAGGAATCTCAGTGTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGCAA 384
QY 903 GTGACCAACATAATCTCAITTTGATGATGAGGAAGATGAGCAGAACTCTCGGGACGTGTTT 962
Db |||
385 GTGACCAACATAATCTCAITTTGATGATGAGGAAGATGAGCAGAACTCTCGGGACATGTTT 444
QY 963 AAAAAAGACCTGGGGAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
Db |||
445 AAAAAAGACCTGGGGAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 504
QY 1023 ATCATGTCGCGCTTGAAGCCCTTTCGGGCTTAACCTCAATGGAA-TCAGAGCAGCAAC 1081
Db |||
505 ATCATGTCGCGCTTGAAGCCCTTTCGGGCTTAACCTCAATGGAAAGTCAAGAGCAGCAAC 564
QY 1082 TCATGAAATTTGATTCCTCTGTTTGAACGGGAGTTCGGGTACCAAGCTTGTATGTG 1141
Db |||
565 TCGTGAATTTGATTCCTCTGTTTGAACGGGAGTTCGGGTACCAAGCTTGTATGTG 624
QY 1142 -AAAGCATCGATGATGAAGATGATGAAGAAAGAAAGATGACGTGTATGGAACCTCATC 1200
Db |||
625 AAAAAAGCATGATGATGAAGATGATGAAGAAAGAAAGATGACGTGTATGGAACCTCATC 684
QY 1201 AGGAAGAGCAGCAGGGCCACTCGAGTCGCCCGAGAA 1239
Db |||
685 AGGACGGAAGCAGGGGCCACTCAGAGTCGCCCGAGAA 723

RESULT 9
BQ057616
LOCUS
DEFINITION BQ057616 1022 bp mRNA linear EST 29-MAR-2002
5', mRNA sequence.
AGENCOURT_7047416 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5813279
ACCESSION BQ057616
VERSION BQ057616.1 GI:19816956
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1022)
NIH-MGC http://mgi.nhl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
```

```
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2065 row: a column: 24
High quality sequence stop: 641.
FEATURES
Location/Qualifiers
source
1..1022
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5813279"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 31.7%; Score 603.8; DB 5; Length 1022;
Best Local Similarity 96.0%; Pred. No. 3.5e-152;
Matches .652; Conservative 0; Mismatches 22; Indels 5; Gaps 3;
QY 544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCCT 603
Db |||
70 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGATTTCCAGTATGCTTCCT 129
QY 604 ACCATGGCAGCAGGTCTGAACCTCATCTCTTTGGCATTAACATCAGCAACAGGATTTG 663
Db |||
130 ACCATGGCAGCAGGTCCGAACCTCATCTCTTTGGCATTAACATGACAAAGGATTTG 189
QY 664 AACGGGAGAGTAAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGAGTCAACGAGAAC 723
Db |||
190 AACGGGAGAGTAAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGAGTCAACGAGAAC 249
QY 724 GTGACCTCTCTGTGAAGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA 783
Db |||
250 GTGA---CCTTGTCTGAAGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA 306
QY 784 GCCTCTCTGCGTCTCCATCTCTCAATCAAACTGAACAGGAGACCGA-CCCTTGCTGTG 842
Db |||
307 GCCTCTCTGCGTCTCCATCTCTCAATCAAACTGAACAGGAGACCGACCCCTTGCCTGTC 366
QY 843 GTGTCAGGAATGTGATGCTGCTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGAA 902
Db |||
367 GTGTCAGGAATGTGATGCTGCTGATGCCAATGCAAAAGGAGCGGAAGAAAGAAAGCAA 426
QY 903 GTGACCAACATAATCTCAITTTGATGATGAGGAAGATGAGCAGAACTCTCGGGACGTGTTT 962
Db |||
427 GTGACCAACATAATCTCAITTTGATGATGAGGAAGATGAGCAGAACTCTCGGGACATGTTT 486
QY 963 AAAAAAGACCTGGGGAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
Db |||
487 AAAAAAGACCTGGGGAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 546
QY 1023 ATCATGTCGCGCTTGAAGCCCTTTCGGGCTTAACCTCAATGGAA-TCAGAGCAGCAAC 1081
Db |||
547 ATCATGTCGCGCTTGAAGCCCTTTCGGGCTTAACCTCAATGGAAAGTCAAGAGCAGCAAC 606
QY 1082 TCATGAAATTTGATTCCTCTGTTTGAACGGGGAGTTCGGGTACCAAGCTTGTATGTG 1141
Db |||
607 TCGTGAATTTGATTCCTCTGTTTGAACGGGGAGTTCGGGTACCAAGCTTGTATGTG 666
QY 1142 AAAAAAGCATGATGAAGATGATGATGAAGAAAGAGATGAGCTGTATGGAACCTCATCA 1201
Db |||
667 AAAAAAGCATGATGAAGATGATGATGAAGAAAGAGATGAGCTGTATGGAACCTCATCA 726
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLCM2065 row: a column: 24  
High quality sequence stop: 641.

FEATURES  
source

1..1022  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5813279"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_99"  
/note="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 31.7%; Score 603.8; DB 5; Length 1022;  
Best Local Similarity 96.0%; Pred. No. 3.5e-152;  
Matches .652; Conservative 0; Mismatches 22; Indels 5; Gaps 3;

```
QY 1202 GGAAGAAAGCACAGGGGCC 1220
Db 727 GCACGAAACACAGGGGGC 745

RESULT 10
BG758425
LOCUS 602712623F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4853164 5',
DEFINITION mRNA sequence.
ACCESSION BG758425
VERSION BG758425.1 GI:14069078
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 843)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1698 row: m column: 05
High quality sequence stop: 835.
Location/Qualifiers
1. 843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4853164"
/tissue_type="Primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 48"
/notes="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 31.5%; Score 600.6; DB 4; Length 843;
Best Local Similarity 96.2%; Pred. No. 2.4e-151;
Matches 659; Conservative 0; Mismatches 19; Indels 7; Gaps 4;

QY 544 AGCACTTTTATGAAGACTGCTTTTGTGATGATGAAGAAGGTCAGATGCTTCCT 603
Db 142 AGTACTTTTATGAAGACTGCTTTTGTGATGATGAAGAAGGTCAGATGCTTCCT 201

QY 604 ACCATGGCAGAGGTCTGAATCCATCTCTTTGCGATTAAATCATCGACAAGGATTTG 663
Db 202 ACCATGGCAGAGGTCTGAATCCATCTCTTTGCGATTAAATCATCGACAAGGATTTG 261

QY 664 AACGGGCAGAGTAAGTTTGTCTCCACCGTTTCAGACCTTTAAAGGAGTCAACGCAAC 723
Db 262 AACGGGCAGAGTAAGTTTGTCTCCACCGTTTCAGACCTTTAAAGGAGTCAACGCAAT 321

QY 724 GTGACCTCTTGTGAAGAGTCCACGCAAGGAGTGAGCAGCTGTTTCAGGGAGATCACA 783
Db 322 GTGA---CCTTGTGAGGAGTCCACGCAAGGAGTGAGCAGCGTTCAGGGAGATCACA 378

QY 784 GCCTCCTCTGCGGTCTCCATCTCATCAAACTGAACAGGAGACCGA-CCCTTCCTGTC 842
```

```
|||||
379 GCCTCCTCTGCGGTCTCCATCTCCATCTCAAACTGAACAGGAGACCGACCCCTTGCCTGTC 438
QY 843 GTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGGAGCGGAAGAAAGAAA 902
Db 439 GTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGGAGCGGAAGAAAGAAA 498
QY 903 GTGACCAACAATACTCTATTGTGATGAGGAAGATGAGCAGAACTCTGGGGAGCTGTTT 962
Db 499 GTGACCAACAATACTCTATTGTGATGAGGAAGATGAGCAGAACTCTGGGGAGCTGTTT 558
QY 963 AAAAAGACACCTGGGGCAGGGAGAGTTCAGAGACAACTCCGACCCCTCTCTGTCAAT 1022
Db 559 AAAAAGACACCTGGGGCAGGGAGAGTTCAGAGACAACTCCGACCCCTCTCTGTCAAT 618
QY 1023 ATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAATCCCAATGGAA-TCAGAGCAGCAAC 1081
Db 619 ATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAATCCCAATGGAAAGTTCAGAGCAGCAAC 678
QY 1082 TCATGG-AAAATTGATTCCTGTCTTTGAAACGGGGAGTTTGGGTACCAAGCTTCATG 1139
Db 679 TCGTGGGAAAATTGGATTCCTGTCTTTGAAACAGGGAGTTTGGGTACCAAGCTTCATG 738
QY 1140 TGAAGAAGCATCGATGATGAAGATGTGATGAAACGAAGATGAGCTGTATGGAACATCAT 1199
Db 739 TGAAGAAGCATCGATGATGAAGATGTGATGAAACGAAGATGAGCTGTATGGAACATCAT 798
QY 1200 CAGGAAGGAAAGCACAGGGGCCACTC 1224
Db 799 CAGGACGGAAGCACAGGGGCCACTC 823

RESULT 11
BQ057008
LOCUS 892 bp mRNA linear EST 29-MAR-2002
DEFINITION AGENCOURT_6769190 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5812156
5', mRNA sequence.
ACCESSION BQ057008
VERSION BQ057008.1 GI:19816348
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 892)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2062 row: c column: 05
High quality sequence stop: 608.
Location/Qualifiers
1. 892
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5812156"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_99"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
```

of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

ORIGIN		Query Match		31.5%; Score 600.4; DB 5; Length 892;		Best Local Similarity		92.1%; Pred. No. 2.8e-151;		Matches 678; Conservative 0; Mismatches 51; Indels 7; Gaps 4;	
QY	544	AGCACATTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCCT	603			QY	544	AGCACATTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCCT	603		
Db	98	AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGTCCAGTATGCTTCCT	157			Db	98	AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGTCCAGTATGCTTCCT	157		
QY	604	ACCATGGCAGCAGTCTGAACCTCCATCTCTTTGCGATTAAACATCGACAAACAGGATTTG	663			QY	604	ACCATGGCAGCAGTCTGAACCTCCATCTCTTTGCGATTAAACATCGACAAACAGGATTTG	663		
Db	158	ACCATGGCAGCAGTCCGAACCTCCATCTCTTTGCGATTAAACATCGACAAACAGGATTTG	217			Db	158	ACCATGGCAGCAGTCCGAACCTCCATCTCTTTGCGATTAAACATCGACAAACAGGATTTG	217		
QY	664	AACGGCAGAGTAAGTTTGTCTCCACGTTTCAGACTCTTAAAGAGTCAACGAGAAC	723			QY	664	AACGGCAGAGTAAGTTTGTCTCCACGTTTCAGACTCTTAAAGAGTCAACGAGAAC	723		
Db	218	AACGGCAGAGTAAGTTTGTCTCCACGTTTCAGACTCTTAAAGAGTCAACGAGAAC	277			Db	218	AACGGCAGAGTAAGTTTGTCTCCACGTTTCAGACTCTTAAAGAGTCAACGAGAAC	277		
QY	724	GTGACCTCTCTGCTGAAGGAGTCCACGCAAGGAGTGCAGGAGTGTTCAGGGAGATCACA	783			QY	724	GTGACCTCTCTGCTGAAGGAGTCCACGCAAGGAGTGCAGGAGTGTTCAGGGAGATCACA	783		
Db	278	GTGAC---CTTGCTGAAGGAGTCCACGCAAGGAGTGCAGGAGTGTTCAGGGAGATCACA	334			Db	278	GTGAC---CTTGCTGAAGGAGTCCACGCAAGGAGTGCAGGAGTGTTCAGGGAGATCACA	334		
QY	784	GCCTCTCTGCCCTCTCCATCTCATCAACCTGAAACAGGAGACCGA-CCCTTGCCTGTC	842			QY	784	GCCTCTCTGCCCTCTCCATCTCATCAACCTGAAACAGGAGACCGA-CCCTTGCCTGTC	842		
Db	335	GCCTCTCTGCCATCTCCATCTCATCAACCTGAAACAGGAGACCGACCCCTTGCCGTC	394			Db	335	GCCTCTCTGCCATCTCCATCTCATCAACCTGAAACAGGAGACCGACCCCTTGCCGTC	394		
QY	843	GTGTCAGGAATGTGAGTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	902			QY	843	GTGTCAGGAATGTGAGTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	902		
Db	395	GTGTCAGGAATGTGAGTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	454			Db	395	GTGTCAGGAATGTGAGTGTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	454		
QY	903	GTGACCAACATATCTCATTTGATGATGAGGAGTGCAGCAACCTCTGGGACGTTT	962			QY	903	GTGACCAACATATCTCATTTGATGATGAGGAGTGCAGCAACCTCTGGGACGTTT	962		
Db	455	GTGACCAACATATCTCATTTGATGATGAGGAGTGCAGCAACCTCTGGGACGATGTT	514			Db	455	GTGACCAACATATCTCATTTGATGATGAGGAGTGCAGCAACCTCTGGGACGATGTT	514		
QY	963	AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT	1022			QY	963	AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT	1022		
Db	515	AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCACTCTCTGTCAAT	574			Db	515	AAAAAGACACCTGGGGCAGGGAGAGCTCAGAGGACAACTCCGACCACTCTCTGTCAAT	574		
QY	1023	ATCATGTCGCTTTGAAAGCCCTTTCGGGCTTAATCCAAATGGAA-TCAGAGCAGCAAC	1081			QY	1023	ATCATGTCGCTTTGAAAGCCCTTTCGGGCTTAATCCAAATGGAA-TCAGAGCAGCAAC	1081		
Db	575	ATCATGTCGCTTTGAAAGCCCTTTCGGGCTTAATCCAAATGGAAATCAGAGCAGCAAC	634			Db	575	ATCATGTCGCTTTGAAAGCCCTTTCGGGCTTAATCCAAATGGAAATCAGAGCAGCAAC	634		
QY	1082	TCATGAAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACCAAGCTTGATGTG	1141			QY	1082	TCATGAAATATGATTCCTCTGTTTGAACGGGAGTTTGGGTACCAAGCTTGATGTG	1141		
Db	635	TCGTGAAATATGATTCCTCTGTTTGAACAGGGAGTTTGGGTACCAAGCTTGATGTG	694			Db	635	TCGTGAAATATGATTCCTCTGTTTGAACAGGGAGTTTGGGTACCAAGCTTGATGTG	694		
QY	1142	AAAGCATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1201			QY	1142	AAAGCATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1201		
Db	695	AAAGCATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	754			Db	695	AAAGCATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	754		
QY	1202	GGAA--GGAAGCAGCAGGGCCACTCGGAGTTCGCGGAGAGCCACTTGAAGGGAACACCT	1259			QY	1202	GGAA--GGAAGCAGCAGGGCCACTCGGAGTTCGCGGAGAGCCACTTGAAGGGAACACCT	1259		
Db	755	GGAGCGAAAGCCAGGGCCACTCCGAGATCCCGCGAGAAACGGAACACACACTTCT	814			Db	755	GGAGCGAAAGCCAGGGCCACTCCGAGATCCCGCGAGAAACGGAACACACACTTCT	814		
QY	1260	GCCTCTCCAGATGCA	1275			QY	1260	GCCTCTCCAGATGCA	1275		
Db	815	GTACACCCAGGCTGGA	830			Db	815	GTACACCCAGGCTGGA	830		

RESULT 12

BM692924  
LOCUS  
DEFINITION  
UI-E-CK1-afk-h-18-0-UI.r1 UI-E-CK1 Homo sapiens cDNA clone  
EST 28-FEB-2002  
UI-E-CK1-afk-h-18-0-UI 5', mRNA sequence.  
ACCESSION  
BM692924  
VERSION  
BM692924.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 609)  
Bonaldo,M.P., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse

FEATURES  
source

Location/Qualifiers  
1..609  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-CK1-afk-h-18-0-UI"  
/tissue\_type="Retina Foveal and Macular"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_libs="UI-E-CK1"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
UI-E-CK1 is a normalized cDNA library containing the  
following tissue(s): Retina Foveal and Macular. The  
library was constructed according to Bonaldo, Lennon and  
Soares, Genome Research, 6:791-806, 1996. First strand  
cDNA synthesis was primed with an oligo-dT primer  
containing a Not I site. Double stranded cDNA was ligated  
to an EcoR I adaptor, digested with Not I, and cloned  
directionally into pT7T3-Pac vector. The oligonucleotide  
used to prime the synthesis of first-strand cDNA contains  
a library tag sequence that is located between the Not I  
site and the (dT)18 tail. The sequence tag for this  
library is GTCC. This library was created for the program,  
Gene Discovery in the Visual System, supported by National  
Eye Institute (NEI)."

ORIGIN

Query Match	30.7%;	Score 586;	DB 4;	Length 609;
Best Local Similarity	99.7%;	Pred. No. 2e-147;		
Matches 608;	Conservative 0;	Mismatches 0;	Indels 2;	Gaps 2;
QY	1215	GGGGCCACTCGGAGTCGCCGAGAGGCACCTGGAGGGAAACACCTGCCCTCTCCAGATGC	1274	
Db	1	GGGGCCACTCGGAGTCGCCGAGAGGCACCTGGAGGGAAACACCTGCCCTCTCCAGATGC	60	
QY	1275	ACAGCTGGGCTCCGCTGAAGGTGCTGCAATGATCTCCGACATCTCTTCCTGTCACTG	1334	
Db	61	ACAGCTGGGCTCCGCTGAAGGTGCTGCAATGATCTCCGACATCTCTTCCTGTCACTG	120	
QY	1335	GGTGGGCTCTACAGCCGAGATGCCCTCGAAGCCTGGAGAACGGGACAGGAC	1394	
Db	121	GGTGGGCTCTACAGCCGAGATGCCCTCGAAGCCTGGAGAACGGGACAGGAC	180	
QY	1395	CAGAGACCACTTCTCCGGATCTCGGTACAGTGTGGAAGCCAGCTCTCCAG	1454	
Db	181	CAGAGACCACTTCTCCGGATCTCGGTACAGTGTGGAAGCCAGCTCTCCAG	240	
QY	1455	GCCACGGAAGTCTCTGAGCAGCCTGTGA-CTTCTGCTCAGTGCAGAGTCCATGACAA	1513	

Db 241 GCCACGGAAGTCTCTGAGCAGCCTGTTACCTTCTGCTCAGTCCAGAGTCCATGACAA 300  
QY 1514 TTAGTGAAGTGGCCAGGCACTGTGGCCATCATGAACAGGAAGATGAGTGGAGGAGG 1573  
Db 301 TTAGTGAAGTGGCCAGGCACTGTGGCCATCATGAACAGGAAGATGAGTGGAGGAGG 360  
QY 1574 AGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGAGCACTCAGCCGCTCCGGC 1633  
Db 361 AGAACAGATCACTGCGAAACCTGCTCGACGGTGAGATGAGCACTCAGCCGCTCCGGC 420  
QY 1634 AAGAGGTGACACCTTTGAAAAGAGGTGGCTGAAACAGAGAGCGGAGGCGCATGAAGG 1693  
Db 421 AAGAGGTGACACCTTTGAAAAGAGGTGGCTGAAACAGAGAGCGGAGGCGCATGAAGG 480  
QY 1694 TCCAGCGCTGCGCAGCTATCTTTGCTATTTTGTGAGGAGATCTTAAACCCACAGTGAGAA 1753  
Db 481 TCCAGCGCTGCGCAGCTATCTTTGCTATTTTGTGAGGAGATCTTAAACCCACAGTGAGAA 540  
QY 1754 CCATGTGTGGAGAAATGGAGGAGAGAGAAATCCACAGTTCTCTGATAGTCTCATTTGA 1813  
Db 541 CCATGTGTGGAGAAATGGAGGAGAGAGAAATCCACAGTTCTCTGATAGTCTCATTTGA 599  
QY 1814 GCTCTCTGAT 1823  
Db 600 GCTCTCTGAT 609

RESULT 13  
BG684262

LOCUS 602635896F1 NIH\_MGC\_48 Homo sapiens cDNA clone linear EST 01-MAY-2001  
DEFINITION mRNA sequence.

ACCESSION BG684262

VERSION BG684262.1 GI:13915659

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 813)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM1619 row: p column: 24

High quality sequence stop: 795.

## FEATURES

source

1. .813  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4763783"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 48"  
/note="Organ: B-cells; Vector: pOTB7; Site: 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >50bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 28.9%; Score 550.8; DB 4; Length 813;  
Best Local Similarity 96.0%; Pred. No. 7.9e-138;  
Matches 651; Conservative 0; Mismatches 17; Indels 10; Gaps 8;  
QY 544 AGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTCCT 603  
Db 142 AGTACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTCCT 200  
QY 604 ACCATGCGCAGCAGTCTGAACTCCATCTCTTTGCGATTAAACATCGAACAAAGGATTTG 663  
Db 201 ACCATGCGCAGCAGTCTGAACTCCATCTCTTTGCGATTAAACATCGAACAAAGGATTTG 260  
QY 664 AACGGGCGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGAGAGTCAACGAGAAC 723  
Db 261 AACGGGCGAGTAAGTTTGTCTCCACCGTTTCAGACCTCTTAAAGAGAGTCAACGAGAAC 320  
QY 724 GTGACCTCTCTTGTGAAGAGAGTCCACGCAAGGAGTGAAGAGTGTTCAGGGAGATCACA 783  
Db 321 GTGA--CCTTGTGAAGAGTCCACGCAAGGAGTGAAGAGTGTTCAGGGAGATCACA 377  
QY 784 GCTCTCTCTGCGTCTCCATCTCAATCAAACTGAAACAGAGACCGA-CCTTTGCGTGTGTC 842  
Db 378 GCTCTCTCTGCGTCTCCATCTCAATCAAACTGAAACAGAGACCGA-CCTTTGCGTGTGTC 437  
QY 843 GTGTCCAGGAATGTCAAGTGTGATGCCAATGCAAAAGAGAGCGGAGAGAAAGAA 902  
Db 438 GTGTCCAGGAATGTCAAGTGTGATGCCAATGCAAAAGAGAGCGGAGAGAAAGAA 497  
QY 903 GTGACCAACAATAATCTCATTTGATGATGAGAAAGTGAAGAGTGTTCGGGAGCTGTTT 962  
Db 498 GTGACCAACAATAATCTCATTTGATGATGAGAAAGTGAAGAGTGTTCGGGAGCTGTTT 557  
QY 963 AAAAGA-CACCTGGGCGAGGAGAGTCTCAGAGCACTCCGACCGTCTCTGTGCAA 1021  
Db 558 AAAAGACCACTGGGCGAGGAGAGTCTCAGAGCACTCCGACCGTCTCTGTGCAA 616  
QY 1022 TATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAATCCCAATGGA-TCAGAGAGCAGAA 1080  
Db 617 TATCATGTCCGCTTTGACAGCCCTTCGGGCGCAAACTCCCAATGGAAGTGTGAGAGCAGAA 676  
QY 1081 CTCATGGAATAAT-GATTCCTCTCTTTGAAACGGGAGTTTGGGTACCAAGACTTGATG 1139  
Db 677 CTCGTGAAATAATGATTCCTCTCTTTGAAACAGGGAG-ITGGGTACCAAGACTTGATG 735  
QY 1140 TGAAGAGCATCGATGATGAAGATGTGATGATAAAGAGATGACGTGTATGGAACATCAT 1199  
Db 736 TGAAGAGCATCGATGATGAAGATGTGATGATAAAGAGATGACGTGTATGGAACATCAT 795  
QY 1200 CAGGAAGGAGCAGCGG 1217  
Db 796 CAGGACGGAAGCAGCGG 813  
RESULT 14  
BQ184171/c  
LOCUS BQ184171  
DEFINITION UI-E-EJ1-ajs-e-23-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone  
ACCESSION BQ184171  
VERSION BQ184171.1 GI:20359722  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 626)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovey  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548

## COMMENT

Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 FORWARD  
POLYA=Yes.

## FEATURES

source

1. .626  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ1-aj5-e-23-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina foveal and Macular, RPE and  
Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-EJ1"  
/notes="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a  
modified polylinker; Site\_1: EcoR I; Site\_2: Not I;  
UI-E-EJ1 is a subtracted cDNA library constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pTT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tags for this library are: fetal eyes,  
AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,  
AATGCGCAT; optic nerve, CCATTAGTG; retina, CCGGCG;  
RPE and Macular, GTCC; RPE and Choroid, ACCTA. This  
library was created for the program, Gene Discovery in the  
Visual System, supported by National Eye Institute (NEI).  
TAG TISSUE=human retina  
TAG\_LIB=UI-E-EJ1  
TAG\_SEQ=CCGGC

## ORIGIN

Query Match 27.4%; Score 522.6; DB 5; Length 626;  
Best Local Similarity 98.9%; Pred. No. 3.3e-130;  
Matches 536; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
362 CAGAGCCCGTGTCTGTACTACGTGAAGAGGTCTCTCAACAGCAGCAGCTCAGCGCT 421  
444 CAGAGCCCGTGTCTGTACTACGTGAAGAGGTCTCTCAACAGCAGCAGCTCAGCGCT 485  
422 TCTACTCCCTGCCGCACATCGCTCAGACGTGGCGCGGGTCCGGCTCGCTCGCTGTG 481  
484 TTTACTCCCTGCCGCACATCGCTCAGACGTGGCGCGGGTCCGGCTCGCTCGCTGTG 425  
482 CCCTCAACGAACACTCCCTGGAGCGGTACTCGACATGCTCTCGGCGGCGCTGCAGGC 541  
424 CCCTCAACGAACACTCCCTGGAGCGGTACTCGACATGCTCTCGGCGGCGCTGCAGGC 365  
542 TGAGCACTTTTATGAAGACTGGTCTTTTGTGTGATGAAGAAAGGTCCAGTATGCTTC 601  
364 TGAGCACTTTTATGAAGACTGGTCTTTTGTGTGATGAAGAAAGGTCCAGTATGCTTC 305  
602 CTACCATGGCAGGCTCTGAATCTCACTCTTTGCGATTATACATCGACAACAGGATT 661  
304 CTACCATGGCAGGCTCTGAATCTCACTCTTTGCGATTATACATCGACAACAGGATT 245

QY 662 TGAACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGA 721  
Db 244 TGAACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGA 185  
QY 722 AGTGACCTCCCTTGTCTGAAGAGGTCCACGCAAGGAGTGAGCAGCTGTTCAGGAGATCA 781  
Db 184 AGTGACCTCCCTTGTCTGAAGAGGTCCACGCAAGGAGTGAGCAGCTGTTCAGGAGATCA 125  
QY 782 CAGCTCTCTCTGCGGTCTCCATCTCATCAAACTCAACAGGAGAGCCGA-CCCTTGGCTG 840  
Db 124 CAGCTCTCTCTGCGGTCTCCATCTCATCAAACTCAACAGGAGAGCCGACCCCTTGGCTG 65  
QY 841 TCGTGTCCAGGAATGTCACTGTGTGATGCCAAATGCAAAAGGAGCGGAAGAAAAAGA 900  
Db 64 TCGTGTCCAGGAATGTCACTGTGTGATGCCAAATGCAAAAGGAGCGGAAGAAAAAGA 5  
QY 901 AA 902  
Db 4 AA 3  
RESULT 15  
BG759681  
LOCUS DEFINITION  
602713374F1 NTH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4853346 5',  
mRNA sequence.  
ACCESSION BG759681  
VERSION BG759681.1 GI:14070334  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 819)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LLCW1699 row: d column: 19  
High quality sequence stop: 791.  
Location/Qualifiers  
1. .819  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4853346"  
/tissue\_type="primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 48"  
/note="Organ: B-cells; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."  
ORIGIN  
Query Match 26.9%; Score 513.2; DB 4; Length 819;  
Best Local Similarity 93.3%; Pred. No. 1.3e-127;  
Matches 603; Conservative 0; Mismatches 33; Indels 10; Gaps 6;  
QY 544 AGCACTTTTATGAGACTGGTCTTTTGTGTGATGAAGAAAGGTCCAGTATGCTTCT 603



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	133.4	7.0	190	4	US-09-513-999C-16002	Sequence 16002, A
2	83.6	4.4	100	4	US-09-513-999C-16212	Sequence 16212, A
3	66.8	3.5	811	4	US-09-270-767-12402	Sequence 12402, A
4	55.6	2.9	7218	1	US-08-232-463-14	Sequence 14, Appl
5	54.8	2.9	36731	4	US-09-949-016-13770	Sequence 13770, A
6	54.4	2.9	6071	4	US-09-949-016-69138	Sequence 69138, A
7	53.2	2.8	476	4	US-09-573-080A-118	Sequence 118, App
8	52.2	2.7	961	4	US-09-976-594-221	Sequence 221, App
9	50	2.6	9811	4	US-09-949-016-12759	Sequence 12759, A
10	46.8	2.5	60276	4	US-09-949-016-15004	Sequence 15004, A
11	46.8	2.5	60338	4	US-09-949-016-15694	Sequence 15694, A
12	46.8	2.5	94133	4	US-09-949-016-11901	Sequence 11901, A
13	46.8	2.5	94133	4	US-09-949-016-12713	Sequence 12713, A
14	46.8	2.5	94135	4	US-09-949-016-15934	Sequence 15934, A
15	46.8	2.5	94135	4	US-09-949-016-15935	Sequence 15935, A
16	46.8	2.5	94135	4	US-09-949-016-15936	Sequence 15936, A
17	46.8	2.5	94135	4	US-09-949-016-15937	Sequence 15937, A
18	46.6	2.4	1182	3	US-09-165-827C-1	Sequence 1, Appl
19	46.6	2.4	1185	3	US-09-165-827C-13	Sequence 13, Appl
20	45.6	2.4	551	4	US-09-621-976-2454	Sequence 2454, Ap
21	45.6	2.4	571	4	US-09-621-976-2453	Sequence 2453, Ap
22	45.4	2.4	1307	4	US-09-506-066E-9	Sequence 9, Appl
23	44.8	2.4	474	4	US-09-621-976-18033	Sequence 18033, A
24	44.8	2.4	2056	4	US-09-498-557-20	Sequence 20, Appl
25	44.6	2.3	23802	4	US-09-949-016-12107	Sequence 12107, A
26	44.6	2.3	23803	4	US-09-949-016-15878	Sequence 15878, A
27	43.8	2.3	76767	4	US-09-949-016-12147	Sequence 12147, A





Qy	1060	CCAATGGAATCAGACCAACAATCATGGAAAATTGATTCCTCGTCTTTTGAAACGGGAGTT	1119
Db	1256	RRR	1197
Qy	1120	TGGGTACCAGAAGCTTGATGTGAAAAGCATCGATGAAAGATGGATGAAAAAACAAGA	1179
Db	1196	RRR	1137
Qy	1180	TGACGTGTATGAAACTCATCAGGAAGAACACAGGGGCCACTCGAGTCCCGCAGAA	1239
Db	1136	RERERRR	1077
Qy	1240	GCCACTGGAAGGGAACA CTGCCTCTCCAGATGCACAGCTGGG	1283
Db	1076	RRRRRRRRRATCGCAAGCTCCCTCGACTGCGAGCAAGCTCGG	1033

## RESULT 5

```

US-09-949-016-13770/c
; Sequence 13770, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13770
; LENGTH: 36731
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13770

```

	Query Match	2.9%;	Score 54.8;	DB 4;	Length 36731;
	Best Local Similarity	63.8%;	Pred. No. 0.00062;		
	Matches 83;	Conservative	0;	Mismatches 47;	Indels 0;
				Gaps	0;
Qy	1768	AATGGGGGAGAGAAATCCACAGTTCTCTGATAGTCTCATTTGAGCTCCTCGATCCAG	1827		
Db	6869	AAAAGTTGAACCAAGACGGAGAAAGATTCTGTGACCTCATTAGAATATCTCGATCCAG	6810		
Qy	1828	TCCTTCTCTGAAGCTGTGTTTCTCTGGACTTTTCATGTATGTGAGCCAAATAAATTGCTTT	1887		
Db	6809	CCTTGCCTGAATGCAGTATACCTCTGGATTCTTAGTCAATGTAGCCAAATGAATCGCGTG	6750		
Qy	1888	CATTCCTTGA	1897		
Db	6749	TTTTCCTTAA	6740		

RESIST 6

RES001 8  
 US-09-949-016-69138/c  
 ; Sequence 69138, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 69138
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-69138

```

### Query Match

Query Match	2.9%	Score 54.4;	DB 4;	Length 601;
Best Local Similarity	63.1%;	Pred. No. 0.0001;		
Matches 82;	Conservative 1;	Mismatches 47;	Indels 0;	Gaps 0;
Qy	1768	AATGGGGGAGAGAGAAATCCAAACAGTTCTCTGATAGTCTCATTTTGAGCTCTCTGGATCCAG	1827	
Db	414	AAAGTTGAACCAAGACGGAAGAAGATTCTGATGACCTCATTTAGATACTCTGGATCCAG	355	
Qy	1828	TCITTTCTCTGAAGCTGTGTTTTCTCTGACATTTTTCATGTATGTGAGCCAAATAAATTGCTTT	1887	
Db	354	CCITGCTGATGCAGTATACCTCTGGATTTCTTAGTCAATGAGCAATGAACGCGCTG	295	
Qy	1888	CATTCCTTGA	1897	
Db	294	TTTTTGCTTAA	285	

## RESULT 7

```

RES001 7
US-09-573-080A-118
? Sequence 118, Application US/09573080A
? Patent No. 6828097
? GENERAL INFORMATION:
? APPLICANT: JOAN, KNOLL
? APPLICANT: ROGAN, PETER
? TITLE OF INVENTION: SINGLE COPY GENOMIC
? FILE REFERENCE: 30307
? CURRENT APPLICATION NUMBER: US/09/573,080A
? CURRENT FILING DATE: 2000-05-16
? NUMBER OF SEQ ID NOS: 479
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 118
? LENGTH: 476
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: repeat_region
? LOCATION: (1)..(476)
? PUBLICATION INFORMATION:
? PUBLICATION INFORMATION:
? AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, J;
? TITLE: Prototypic sequences for human repeat
? JOURNAL: Journal of Molecular Evolution
? VOLUME: 35
? ISSUE: 4
? PAGES: 286-291
? DATE: 1992-10-
? DATABASE ACCESSION NUMBER: Database of repeats
? DATABASE ENTRY DATE:
? DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-118

```

### Query Match

	Query Match	2.8%	Score 53.2;	DB 4;	Length 476;
	Best Local Similarity	69.5%;	Pred. No. 0.00019;		
	Matches 89;	Conservative 0;	Mismatches 33;	Indels 6;	Gaps 1;
Qy	1776	GAGAGAGAAATCAAACAGTTTCCTGATAGTCTCAATTTGAGCTCTCTGGATCCAGTCTTTCCT	1835		
Db	313	GAGATGGAGAGACAGAGTTCTCTGATGACATCATTTTGAGCACCTGGATCCAGGCATGCGCT	372		
Qy	1836	GAGG-----CTGTGTTTTCTCTGGACTTTTCATGATGATGGAGCCAAATAAATGCTTTTCA	1889		
Db	373	GAAGCCAGAGTCTATGTCACCCCTGGACTTTTTCAGTTACATGAGCCAAATAAATCTCTTTT	432		

```
QY 1890 TTCCTTGA 1897
|||
Db 433 TTGCTTAA 440
|||

RESULT 8
US-09-976-594-221
; Sequence 221, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 221
; LENGTH: 961
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 330927.1
US-09-976-594-221

Query Match 2.7%; Score 52.2; DB 4; Length 961;
Best Local Similarity 71.1%; Pred. No. 0.00052;
Matches 69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1793 GTTCTGATGTCATTTAGCTCCTGGATCCAGTCTTTCTCTGAAGCTGTGTTCTCT 1852
|||
Db 861 GTCTTGTGTCATCTTTTGGCCCACTAGACCAAGCTTTTACCTGAAGCAGACTACCTCA 920
|||

QY 1853 GGACTTTTTCATGTATGTGAGCCCAATAAATTCCTTTCA 1889
|||
Db 921 GAACCTTTTTCAGTATGTGAGCCCAATAAATTCCTTTCA 957
|||

RESULT 9
US-09-949-016-12759
; Sequence 12759, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12759
; LENGTH: 9811
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12759

Query Match 2.6%; Score 50; DB 4; Length 9811;
Best Local Similarity 61.5%; Pred. No. 0.0067;
Matches 80; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 1768 AATGAGGGAGAGAGAAATCCACAGTTCCTGATGTCTATTTGAGCTCCTGGATCCAG 1827
|||
```

```
Db 3752 AAGAAAGAGAGTGTAGTGAGTGACCATTTCTTTATTCCATCTCTTTGAGGCCCTACACCTAA 3811
|||
QY 1828 TCCTTCTCTGAAGCTGTGTTCCCTCTGACACTTTTCATGATGTAGGCCAATAAATGCTTT 1887
|||
Db 3812 TCATTCAATGAAGCTATATACCACTCCAGCCTTTTCAATTACTTGAGCCAATCAATTCCTT 3871
|||

QY 1888 CATTCTTTGA 1897
|||
Db 3872 TAATAGTTAA 3881
|||

RESULT 10
US-09-949-016-15004/c
; Sequence 15004, Application US/09949016
; Patent No. 8812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15004
; LENGTH: 60276
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(60276)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15004

Query Match 2.5%; Score 46.8; DB 4; Length 60276;
Best Local Similarity 73.2%; Pred. No. 0.13;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1806 TCATTGAGCTCCTGGATCCAGTCTTTCTCTGAAGCTGTGTTCTCTGAGCTTTTCATGT 1865
|||
Db 10119 TCTTAGAAGCACCCTGGATCCAGCCCTCCCTGAGGCCATCTTACCCCTGAGCTTCTCAGTT 10060
|||

QY 1866 ATGTGAGCCCAATAAATTCGCTTT 1887
|||
Db 10059 GTGTGAGCCCAAGAATTCGCTTT 10038
|||

RESULT 11
US-09-949-016-15694/c
; Sequence 15694, Application US/09949016
; Patent No. 8812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15694
```

```
; SEQ ID NO 15694
; LENGTH: 60338
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(60338)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15694

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 60338;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1806 TCATTGAGCTCTGATCAGTCCTTCTGAGCTGTTCTCTGAGCTTTTCATCT 1865
Db 10181 TCTTAGAGCACCCTGATCAGCCCTGCTGAGCCATCTACCCCTGAGCTTCTCAGTT 10122

QY 1866 ATCTGAGCCCAATAAATGCTTT 1887
Db 10121 GTGTGAGCCCAAGAATGCTTT 10100

RESULT 12
US-09-949-016-11901/c
; Sequence 11901, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11901
; LENGTH: 94133
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11901

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 94133;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1792 AGTTCCTGATAGTCTCATTTGAGCTCCTGATCCAGTCTTTCTGAGCTGTTTCTCCTC 1851
Db 40227 AGAGCCTGAGACACAGTTTAAAGCATTTGGAGTCAGCTATGCTGAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGATGTGAGCCCAATAAATGCTTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCCAATAAATATTTTA 40130

RESULT 13
US-09-949-016-12713/c
; Sequence 12713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12713
; LENGTH: 94133
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12713

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 94133;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1792 AGTTCCTGATAGTCTCATTTGAGCTCCTGATCCAGTCTTTCTGAGCTGTTTCTCCTC 1851
Db 40227 AGAGCCTGAGACACAGTTTAAAGCATTTGGAGTCAGCTATGCTGAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGATGTGAGCCCAATAAATGCTTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCCAATAAATATTTTA 40130

RESULT 14
US-09-949-016-15934/c
; Sequence 15934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15934
; LENGTH: 94135
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15934

Query Match
Best Local Similarity 2.5%; Score 46.8; DB 4; Length 94135;
Matches 66; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1792 AGTTCCTGATAGTCTCATTTGAGCTCCTGATCCAGTCTTTCTGAGCTGTTTCTCCTC 1851
Db 40227 AGAGCCTGAGACACAGTTTAAAGCATTTGGAGTCAGCTATGCTGAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGATGTGAGCCCAATAAATGCTTTCA 1889
Db 40167 TAGACTTTTCAGTTACATGAGCCCAATAAATATTTTA 40130

RESULT 15
US-09-949-016-15935/c
; Sequence 15935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15935
; LENGTH: 94135
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15935
```

us-09-155-676b-1.rni

Mon Jun 13 13:27:09 2005

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15935
; LENGTH: 94135
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15935
```

```
Query Match      2.5%; Score 46.8; DB 4; Length 94135;
Best Local Similarity 67.3%; Pred. No. 0.16; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 32;

QY 1792 AGTTCCCTGATAGTCTCTCATTGAGCTCCTGGATCCAGTCTTTCTGAAAGCTGTTCCCTC 1851
Db 40227 AGAGCTCGAAGAACACAGTTTAAGCATTTGGAGTCAGCTATGCTGAAAGCCATGCTCCCTC 40168

QY 1852 TGGACTTTTCATGTTATGTGAGCCATAAATTGCTTCA 1889
Db 40167 TAGACTTTTCAGTTTACATGAGCCATAAAATATTTT 40130
```

Search completed: June 10, 2005, 20:46:32  
Job time : 213.376 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 4767.2 Seconds  
(without alignments)  
16321.674 Million cell updates/sec

Title: US-09-155-676B-1

Perfect score: 1906

Sequence: 1 catgggtcagcggtggcg.....tcattcttgaaaaaaa 1906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/PCTUS1 COMB.seq:\*

2: /cgn2\_6/ptodata/1/pna/PCTUS2 COMB.seq:\*

3: /cgn2\_6/ptodata/1/pna/PCTUS3 COMB.seq:\*

4: /cgn2\_6/ptodata/1/pna/US06 COMB.seq:\*

5: /cgn2\_6/ptodata/1/pna/US07 COMB.seq:\*

6: /cgn2\_6/ptodata/1/pna/US08 COMB.seq:\*

7: /cgn2\_6/ptodata/1/pna/US081 COMB.seq:\*

8: /cgn2\_6/ptodata/1/pna/US082 COMB.seq:\*

9: /cgn2\_6/ptodata/1/pna/US083 COMB.seq:\*

10: /cgn2\_6/ptodata/1/pna/US084 COMB.seq:\*

11: /cgn2\_6/ptodata/1/pna/US085 COMB.seq:\*

12: /cgn2\_6/ptodata/1/pna/US086 COMB.seq:\*

13: /cgn2\_6/ptodata/1/pna/US087 COMB.seq:\*

14: /cgn2\_6/ptodata/1/pna/US088 COMB.seq:\*

15: /cgn2\_6/ptodata/1/pna/US089 COMB.seq:\*

16: /cgn2\_6/ptodata/1/pna/US090 COMB.seq:\*

17: /cgn2\_6/ptodata/1/pna/US091 COMB.seq:\*

18: /cgn2\_6/ptodata/1/pna/US092A COMB.seq:\*

19: /cgn2\_6/ptodata/1/pna/US092B COMB.seq:\*

20: /cgn2\_6/ptodata/1/pna/US093A COMB.seq:\*

21: /cgn2\_6/ptodata/1/pna/US093B COMB.seq:\*

22: /cgn2\_6/ptodata/1/pna/US094 COMB.seq:\*

23: /cgn2\_6/ptodata/1/pna/US095A COMB.seq:\*

24: /cgn2\_6/ptodata/1/pna/US095B COMB.seq:\*

25: /cgn2\_6/ptodata/1/pna/US095C COMB.seq:\*

26: /cgn2\_6/ptodata/1/pna/US095D COMB.seq:\*

27: /cgn2\_6/ptodata/1/pna/US096A COMB.seq:\*

28: /cgn2\_6/ptodata/1/pna/US096B COMB.seq:\*

29: /cgn2\_6/ptodata/1/pna/US096C COMB.seq:\*

30: /cgn2\_6/ptodata/1/pna/US096D COMB.seq:\*

31: /cgn2\_6/ptodata/1/pna/US096E COMB.seq:\*

32: /cgn2\_6/ptodata/1/pna/US097A COMB.seq:\*

33: /cgn2\_6/ptodata/1/pna/US097B COMB.seq:\*

34: /cgn2\_6/ptodata/1/pna/US097C COMB.seq:\*

35: /cgn2\_6/ptodata/1/pna/US098A COMB.seq:\*

36: /cgn2\_6/ptodata/1/pna/US098B COMB.seq:\*

37: /cgn2\_6/ptodata/1/pna/US098C COMB.seq:\*

38: /cgn2\_6/ptodata/1/pna/US098D COMB.seq:\*

39: /cgn2\_6/ptodata/1/pna/US099A COMB.seq:\*

40: /cgn2\_6/ptodata/1/pna/US099B COMB.seq:\*

41: /cgn2\_6/ptodata/1/pna/US099C COMB.seq:\*

42: /cgn2\_6/ptodata/1/pna/US099D COMB.seq:\*

43: /cgn2\_6/ptodata/1/pna/US099E COMB.seq:\*

44: /cgn2\_6/ptodata/1/pna/US099F COMB.seq:\*

45: /cgn2\_6/ptodata/1/pna/US100A COMB.seq:\*

46: /cgn2\_6/ptodata/1/pna/US100B COMB.seq:\*

47: /cgn2\_6/ptodata/1/pna/US100C COMB.seq:\*

48: /cgn2\_6/ptodata/1/pna/US101A COMB.seq:\*

49: /cgn2\_6/ptodata/1/pna/US101B COMB.seq:\*

50: /cgn2\_6/ptodata/1/pna/US102A COMB.seq:\*

51: /cgn2\_6/ptodata/1/pna/US102B COMB.seq:\*

52: /cgn2\_6/ptodata/1/pna/US103A COMB.seq:\*

53: /cgn2\_6/ptodata/1/pna/US103B COMB.seq:\*

54: /cgn2\_6/ptodata/1/pna/US104A COMB.seq:\*

55: /cgn2\_6/ptodata/1/pna/US104B COMB.seq:\*

56: /cgn2\_6/ptodata/1/pna/US105A COMB.seq:\*

57: /cgn2\_6/ptodata/1/pna/US105B COMB.seq:\*

58: /cgn2\_6/ptodata/1/pna/US106A COMB.seq:\*

59: /cgn2\_6/ptodata/1/pna/US107A COMB.seq:\*

60: /cgn2\_6/ptodata/1/pna/US107B COMB.seq:\*

61: /cgn2\_6/ptodata/1/pna/US107C COMB.seq:\*

62: /cgn2\_6/ptodata/1/pna/US107D COMB.seq:\*

63: /cgn2\_6/ptodata/1/pna/US108A COMB.seq:\*

64: /cgn2\_6/ptodata/1/pna/US108B COMB.seq:\*

65: /cgn2\_6/ptodata/1/pna/US109A COMB.seq:\*

66: /cgn2\_6/ptodata/1/pna/US109B COMB.seq:\*

67: /cgn2\_6/ptodata/1/pna/US109C COMB.seq:\*

68: /cgn2\_6/ptodata/1/pna/US110 COMB.seq:\*

69: /cgn2\_6/ptodata/1/pna/US6000 COMB.seq:\*

70: /cgn2\_6/ptodata/1/pna/US6001 COMB.seq:\*

71: /cgn2\_6/ptodata/1/pna/US6002 COMB.seq:\*

72: /cgn2\_6/ptodata/1/pna/US6003 COMB.seq:\*

73: /cgn2\_6/ptodata/1/pna/US6004 COMB.seq:\*

74: /cgn2\_6/ptodata/1/pna/US6005 COMB.seq:\*

75: /cgn2\_6/ptodata/1/pna/US6006 COMB.seq:\*

76: /cgn2\_6/ptodata/1/pna/US6007 COMB.seq:\*

77: /cgn2\_6/ptodata/1/pna/US6008 COMB.seq:\*

78: /cgn2\_6/ptodata/1/pna/US6009 COMB.seq:\*

79: /cgn2\_6/ptodata/1/pna/US6010 COMB.seq:\*

80: /cgn2\_6/ptodata/1/pna/US6011 COMB.seq:\*

81: /cgn2\_6/ptodata/1/pna/US6012 COMB.seq:\*

82: /cgn2\_6/ptodata/1/pna/US6013 COMB.seq:\*

83: /cgn2\_6/ptodata/1/pna/US6014 COMB.seq:\*

84: /cgn2\_6/ptodata/1/pna/US6015 COMB.seq:\*

85: /cgn2\_6/ptodata/1/pna/US6016 COMB.seq:\*

86: /cgn2\_6/ptodata/1/pna/US6017 COMB.seq:\*

87: /cgn2\_6/ptodata/1/pna/US6018 COMB.seq:\*

88: /cgn2\_6/ptodata/1/pna/US6019 COMB.seq:\*

89: /cgn2\_6/ptodata/1/pna/US6020 COMB.seq:\*

90: /cgn2\_6/ptodata/1/pna/US6021 COMB.seq:\*

91: /cgn2\_6/ptodata/1/pna/US6022 COMB.seq:\*

92: /cgn2\_6/ptodata/1/pna/US6023A COMB.seq:\*

93: /cgn2\_6/ptodata/1/pna/US6023B COMB.seq:\*

94: /cgn2\_6/ptodata/1/pna/US6024 COMB.seq:\*

95: /cgn2\_6/ptodata/1/pna/US6025 COMB.seq:\*

96: /cgn2\_6/ptodata/1/pna/US6026 COMB.seq:\*

97: /cgn2\_6/ptodata/1/pna/US6027 COMB.seq:\*

98: /cgn2\_6/ptodata/1/pna/US6028 COMB.seq:\*

99: /cgn2\_6/ptodata/1/pna/US6029 COMB.seq:\*

100: /cgn2\_6/ptodata/1/pna/US6030 COMB.seq:\*

101: /cgn2\_6/ptodata/1/pna/US6031 COMB.seq:\*

102: /cgn2\_6/ptodata/1/pna/US6032 COMB.seq:\*

103: /cgn2\_6/ptodata/1/pna/US6033 COMB.seq:\*

104: /cgn2\_6/ptodata/1/pna/US6034 COMB.seq:\*

105: /cgn2\_6/ptodata/1/pna/US6035 COMB.seq:\*

106: /cgn2\_6/ptodata/1/pna/US6036 COMB.seq:\*

107: /cgn2\_6/ptodata/1/pna/US6037 COMB.seq:\*

108: /cgn2\_6/ptodata/1/pna/US6038 COMB.seq:\*

109: /cgn2\_6/ptodata/1/pna/US6039 COMB.seq:\*

110: /cgn2\_6/ptodata/1/pna/US6040 COMB.seq:\*

111: /cgn2\_6/ptodata/1/pna/US6041 COMB.seq:\*

112: /cgn2\_6/ptodata/1/pna/US6042 COMB.seq:\*

113: /cgn2\_6/ptodata/1/pna/US6043 COMB.seq:\*

114: /cgn2\_6/ptodata/1/pna/US6044 COMB.seq:\*

115: /cgn2\_6/ptodata/1/pna/US6045 COMB.seq:\*

116: /cgn2\_6/ptodata/1/pna/US6046 COMB.seq:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1899	99.6	1906	17	US-09-155-676-1	Sequence 1, Appli
2	1899	99.6	1906	17	US-09-155-676A-1	Sequence 1, Appli
3	1742	91.4	1782	47	US-10-070-255-4	Sequence 4, Appli
4	1555	81.6	2873	47	US-10-070-255-5	Sequence 6, Appli
5	1555	81.6	3139	47	US-10-070-255-5	Sequence 5, Appli
6	1138.2	59.7	2248	28	US-09-629-469A-14016	Sequence 14016, A
7	1138.2	59.7	2248	65	US-10-917-503-14016	Sequence 4362, Ap
8	1092.4	57.3	1516	22	US-09-471-275-4362	Sequence 9141, Ap
9	1092.4	57.3	1516	26	US-09-488-725B-9141	Sequence 9141, Ap
10	1092.4	57.3	1516	26	US-09-552-317-9141	Sequence 9141, Ap
11	883.6	46.4	1137	33	US-09-758-472-1043	Sequence 1043, Ap
12	883.6	46.4	1137	50	US-10-238-926-1043	Sequence 1043, Ap
13	792.8	41.6	6045	28	US-09-629-469A-18174	Sequence 18174, A
14	792.8	41.6	6045	65	US-10-917-503-18174	Sequence 18174, A
15	777.4	40.8	1630	90	US-60-212-356-340	Sequence 340, App
16	774.2	40.6	1642	90	US-60-213-356-373	Sequence 373, App
17	765.6	40.2	814	49	US-10-170-235-20448	Sequence 20448, A
18	738.8	38.8	1147	48	US-10-144-771-19120	Sequence 19120, A
19	738.8	38.8	1147	106	US-60-360-207-19120	Sequence 19120, A
20	719.8	37.8	834	21	US-09-396-087-4235	Sequence 4235, Ap
21	672.2	35.3	730	28	US-09-629-469A-1808	Sequence 1808, Ap
22	672.2	35.3	730	65	US-10-917-503-1808	Sequence 1808, Ap
23	565.4	29.7	575	25	US-09-540-212A-52972	Sequence 52972, A
24	554.4	29.1	575	97	US-60-278-561-628	Sequence 628, App
25	535	28.1	542	21	US-09-540-229-98596	Sequence 98596, A
26	456.4	23.9	512	25	US-09-396-087-160	Sequence 160, App
27	446.2	23.4	554	49	US-10-170-235-23933	Sequence 23933, A
28	440.4	22.7	1667	87	US-60-185-213-1385	Sequence 1385, Ap
29	433.6	22.9	506	25	US-09-540-212A-44800	Sequence 44800, A
30	421.8	22.1	437	21	US-09-362-510-18241	Sequence 18241, A
31	421.8	22.1	437	21	US-09-362-510A-18241	Sequence 18241, A
32	421.8	22.1	437	39	US-09-904-013-18241	Sequence 18241, A
33	417.4	21.9	2075	51	US-60-229-525-567	Sequence 567, App
34	412.4	21.6	422	23	US-09-528-409-108313	Sequence 108313, A
35	412.4	21.6	422	40	US-09-933-524A-108313	Sequence 108313, A
36	412.4	21.6	422	40	US-09-933-524A-108313	Sequence 108313, A
37	404.6	21.2	455	23	US-09-528-409-108118	Sequence 108118, A
38	404.6	21.2	455	40	US-09-933-524-108118	Sequence 108118, A
39	404.6	21.2	455	49	US-09-933-524A-108118	Sequence 108118, A
40	393	20.6	1547	49	US-10-170-235-6644	Sequence 6644, Ap
41	388.2	20.4	406	19	US-09-287-618-7448	Sequence 7448, Ap
42	382	20.0	422	23	US-09-528-409-78923	Sequence 78923, A
43	382	20.0	422	40	US-09-933-524-78923	Sequence 78923, A
44	382	20.0	422	40	US-09-933-524A-78923	Sequence 78923, A
45	374.4	19.6	391	97	US-60-378-561-827	Sequence 827, App

## ALIGNMENTS

ADDRESS: BROWDY AND NEIMAR, P.L.L.C.  
STREET. 624 Ninth Street. N.W., Suite 300

STREET: 021 ARDEN STREET, NEW YORK, NY 10017  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20001

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SERIAL: Patent In Release #1.0.
VERSION: #1.30

```

```

; SOFTWARE: FACILCIN RELEASE W1.0.7
;
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676
; FILING DATE: 04-JAN-1999
;

```

FILING DATE: 04-04-1997  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/IL97/00117  
 FILING DATE: 01-APR-1997

FILING DATE: 02-APR-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: IL 117800  
 FILING DATE: 02-APR-1996

PRIOR APPLICATION DATA: IL 119133  
APPLICATION NUMBER: IL 119133  
FILING DATE: 26-AUG-1996

ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: WALLACH=21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 1906 base pairs
;
; TYPE: nucleic acid
;
; STRANDEDNESS: single
;
;

```

```

;      TOPOLOGY: linear
;      MOLECULE TYPE:  CDNA
US-09-155-676-1

```

Query Match	99.6%;	Score 1899;	DB 17;	Length 1906;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1906. Conservative	0.	Mismatches	0.	Indels 0;

Matches	1906;	Conservative	0;	Mismatches	0;	Inserts	0;	Deletions
Qy	1	CATTGGCTCACGCGGTGGCGCGCTCTAGAAATAGTGGAATCCCCGGCTGCAGGAATCG	60					

[illegible][illegible]

	Qy	Db	
121	AGCGGATCNCNGAACNATGACAAAAGCAATTTCTGCTGGAGCGCATGCTGGATGCACTG	121	AGCGGATCNCNGAACNATGACAAAAGCAATTTCTGCTGGAGCGCATGCTGGATGCACTG
180		180	

181 AAA CAG TCC AGAT CCG CTT TNG AGG GGA AAG GAG ATT G CCT CGG ATT T CCG A CAG CAGG 240

QY	241	GTCACTCTCTGTGTGCGCCAGTTTGAAGCCGTCTCTGCAGCATGCGCTTGAAGAGGAGTCGA	300
DB	241	GTCACTCTCTGTGTGCGCCAGTTTGAAGCCGTCTCTGCAGCATGCGCTTGAAGAGGAGTCGA	300
QY	301	GGATTGGCACTCACAGCGGCAGCGATCAAGCAGCAGCGCGGCTTTGCCAGMAAACCGAA	360
DB	301	GGATTGGCACTCACAGCGGCAGCGATCAAGCAGCAGCGCGGCTTTGCCAGMAAACCGAA	360
QY	361	ACAGAGCCCGTGTTCTGTGTACTACGTGAAGGAGGTCTCTCAACAGCAGCAGCTGCAGCGC	420
DB	361	ACAGAGCCCGTGTTCTGTGTACTACGTGAAGGAGGTCTCTCAACAGCAGCAGCTGCAGCGC	420
QY	421	TTCTACTCCCTCGCGCCACATCGCCTCAGACGTGGGCGGGGTTCGCGCTGCGCTGT	480
DB	421	TTCTACTCCCTCGCGCCACATCGCCTCAGACGTGGGCGGGGTTCGCGCTGCGCTGT	480
QY	481	GCCCTCAACGAACTCCCTCGGAGGCTTACCTGCACATGCTCTCGGCGGACCGCTGCAGG	540
DB	481	GCCCTCAACGAACTCCCTCGGAGGCTTACCTGCACATGCTCTCGGCGGACCGCTGCAGG	540
QY	541	CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT	600
DB	541	CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT	600
QY	601	CCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTTGGATTAACATCGACAACAGGAT	660
DB	601	CCTACCATGGCAGCAGGTCTGAACTCCATACTCTTTTGGATTAACATCGACAACAGGAT	660
QY	661	TTGAAACGGGCAGAGTAAAGTTTGGCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAG	720
DB	661	TTGAAACGGGCAGAGTAAAGTTTGGCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAG	720
QY	721	AACGTGACCTCTTGTGTAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATC	780
DB	721	AACGTGACCTCTTGTGTAAGGAGTCCACGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATC	780
QY	781	ACAGCCTCCTCTGCGGCTCTCATCTCATCAAACTGAAAGGAGACCGACCTTGCCTG	840
DB	781	ACAGCCTCCTCTGCGGCTCTCATCTCATCAAACTGAAAGGAGACCGACCTTGCCTG	840
QY	841	TCGTGTCAGGAATGTGAGTCTCTGATGCCAAATCCAAAAGGAGCGGAAGAAAAGA	900
DB	841	TCGTGTCAGGAATGTGAGTCTCTGATGCCAAATCCAAAAGGAGCGGAAGAAAAGA	900
QY	901	AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAAACTCTGGGGACGTGT	960
DB	901	AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAAACTCTGGGGACGTGT	960
QY	961	TTAAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA	1020
DB	961	TTAAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCA	1020
QY	1021	ATATCATGTCCGCTTTGAAAAGCCCTTCGGGCCCTAACTCCAATGGAAATCAGAGCAGCAA	1080
DB	1021	ATATCATGTCCGCTTTGAAAAGCCCTTCGGGCCCTAACTCCAATGGAAATCAGAGCAGCAA	1080
QY	1081	CTCATGGAAAAATTGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGACTTTGATGT	1140
DB	1081	CTCATGGAAAAATTGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGACTTTGATGT	1140
QY	1141	GAAAAGCATCGATGATGAAGATGTGGATGAAAAAGGAGATGACGTGTATGGAATCTCATC	1200
DB	1141	GAAAAGCATCGATGATGAAGATGTGGATGAAAAAGGAGATGACGTGTATGGAATCTCATC	1200
QY	1201	AGGAAGGAAGCACAGGGGCCACTCGGAGTCCCGCGAGAGCCACTGGAAGGGAACACCTG	1260
DB	1201	AGGAAGGAAGCACAGGGGCCACTCGGAGTCCCGCGAGAGCCACTGGAAGGGAACACCTG	1260
QY	1261	CCTCTCCAGATGCACAGCTGGGCTTCGCTGAAGGTCTGCACAAATGACTCCGACATCCT	1320
DB	1261	CCTCTCCAGATGCACAGCTGGGCTTCGCTGAAGGTCTGCACAAATGACTCCGACATCCT	1320
QY	1321	CTTCCCTGTCTGAGTGGGCTCTCTACAGCCACAGCAGATGCCCCCTTCGGAAGCCTCGA	1380

RESULT 2  
US-09-155-676A-1  
Sequence 1, Application US/09155676A  
GENERAL INFORMATION:  
APPLICANT: WALLACH, David  
APPLICANT: MALININ, Nikolai  
APPLICANT: BOLDIN, Mark  
APPLICANT: KOVALENKO, Andrei  
APPLICANT: METT, Igor  
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
STREET: 624 Ninth Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/155,676A  
FILING DATE: 04-APR-1999  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IL97/00117  
FILING DATE: 01-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 117800

FILING DATE: 02-APR-1996  
PRIOR APPLICATION DATA: IL 119133  
FILING DATE: 26-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: WALLACH-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1906 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-155-676A-1

Query Match 99.6%; Score 1899; DB 17; Length 1906;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1906; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CATTGGGTACGCGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCAGGAATTCG	60
DB	1	CATTGGGTACGCGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCAGGAATTCG	60
QY	61	ATTCCAGGCCACGAAGCGCGGCGCGCANGCACCGGCCGGGAGNAGGCCATG	120
DB	61	ATTCCAGGCCACGAAGCGCGGCGCGCANGCACCGGCCGGGAGNAGGCCATG	120
QY	121	AGCGGATCNCNGAANCATGACAAAAAGACAAATTTCTGTGGAGCGACTCTGGATGCAATG	180
DB	121	AGCGGATCNCNGAANCATGACAAAAAGACAAATTTCTGTGGAGCGACTCTGGATGCAATG	180
QY	181	AAACAGTGCAGATCCGCTTTNGAGGAGAGAGATTTGCTCGGATTCGACAGCAGG	240
DB	181	AAACAGTGCAGATCCGCTTTNGAGGAGAGAGATTTGCTCGGATTCGACAGCAGG	240
QY	241	GTCACTGTCTGTGTGCCAGTTTGAAGCCGTCTCGACAGTGGCTTGAAGAGAGTGA	300
DB	241	GTCACTGTCTGTGTGCCAGTTTGAAGCCGTCTCGACAGTGGCTTGAAGAGAGTGA	300
QY	301	GGATTGGCACTACAGCGCAGGATCAAGCAGGCGGGCTTTGCCAGCAAAACCGAA	360
DB	301	GGATTGGCACTACAGCGCAGGATCAAGCAGGCGGGCTTTGCCAGCAAAACCGAA	360
QY	361	ACAGAGCCGTGTTCTGGTACTACGTGAAGGAGTCCCTCAACAGCACAGCTGCAGCGC	420
DB	361	ACAGAGCCGTGTTCTGGTACTACGTGAAGGAGTCCCTCAACAGCACAGCTGCAGCGC	420
QY	421	TTCTACTCTCCGCGCACATCGCCTCAGAGTGGGCGGGGTGCGGCTGGCTGGCTGT	480
DB	421	TTCTACTCTCCGCGCACATCGCCTCAGAGTGGGCGGGGTGCGGCTGGCTGGCTGT	480
QY	481	GCCTCAACGAAACACTCCCTGGAGCGCTACCTGCACATGCTCTGGCCGACCGCTGCAGG	540
DB	481	GCCTCAACGAAACACTCCCTGGAGCGCTACCTGCACATGCTCTGGCCGACCGCTGCAGG	540
QY	541	CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT	600
DB	541	CTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGGATGAAGAAAGGTCCAGTATGCTT	600
QY	601	CTTACCATGGCAGAGTCTGAATCCATCTCTTTGCGATTAACTCGACACAGGAT	660
DB	601	CTTACCATGGCAGAGTCTGAATCCATCTCTTTGCGATTAACTCGACACAGGAT	660
QY	661	TTGAACGGGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAG	720
DB	661	TTGAACGGGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGCGAG	720
QY	721	AACGTGACCTCTTGTCTGAAGAGTCCACGCAAGGAGTGAAGCCCTGTTTCAGGGAGATC	780

DB	721	AACGTGACCTCTTGTCTGAAGAGTCCACGCAAGAGTGAAGCCCTGTTTCAGGGAGATC	780
QY	781	ACAGCCTCTCTGCGGTCTCCATCTCTATCAAACTGAACAGGAGACCGCCCTTGGCTG	840
DB	781	ACAGCCTCTCTGCGGTCTCCATCTCTATCAAACTGAACAGGAGACCGCCCTTGGCTG	840
QY	841	TGCTGTCCAGGAATGTCTAGTCTGATGCAAAATGCAAAAGAGCGGAGAGAAAAAGA	900
DB	841	TCGTGTCCAGGAATGTCTAGTCTGATGCAAAATGCAAAAGAGCGGAGAGAAAAAGA	900
QY	901	AAAGTCAACCAATAATCTCATTTGATGATGAGGAAGATGAGAGAACTCTTGGGACGTGT	960
DB	901	AAAGTCAACCAATAATCTCATTTGATGATGAGGAAGATGAGAGAACTCTTGGGACGTGT	960
QY	961	TTAAAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTCTGCA	1020
DB	961	TTAAAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTCTGCA	1020
QY	1021	ATATCATGTCCGCTTTGAAAGCCCTTTCCGGCCCTTAACCTCAATGGAATCAGAGCAGAA	1080
DB	1021	ATATCATGTCCGCTTTGAAAGCCCTTTCCGGCCCTTAACCTCAATGGAATCAGAGCAGAA	1080
QY	1081	CTCATGGAATAATGATTCCTCTGTGAAACGGGAGTTTGGGTACCAAGACTTGATGT	1140
DB	1081	CTCATGGAATAATGATTCCTCTGTGAAACGGGAGTTTGGGTACCAAGACTTGATGT	1140
QY	1141	GAAAGCATCGATGATGAAGATGTGATGAAACGAGATGACGTGTATGTAATCTCATC	1200
DB	1141	GAAAGCATCGATGATGAAGATGTGATGAAACGAGATGACGTGTATGTAATCTCATC	1200
QY	1201	AGGAAGAAAGCACAGGGGCCACTCGAGTCCGCCGAGAAAGCCACTGGAAGGAAACACCTG	1260
DB	1201	AGGAAGAAAGCACAGGGGCCACTCGAGTCCGCCGAGAAAGCCACTGGAAGGAAACACCTG	1260
QY	1261	CCTCTCCAGATGCAAGCTGGGCTCCGCTGAAGTGTGCTGCAAACTCTCCGACATCTCT	1320
DB	1261	CCTCTCCAGATGCAAGCTGGGCTCCGCTGAAGTGTGCTGCAAACTCTCCGACATCTCT	1320
QY	1321	CTTCCCTGTCTGAGTGGGCTCTCTACAGCCAGCAGATGCCCCCTCGGAAGCCCTGGA	1380
DB	1321	CTTCCCTGTCTGAGTGGGCTCTCTACAGCCAGCAGATGCCCCCTCGGAAGCCCTGGA	1380
QY	1381	GAAAGGAGACAGGACAGAGACCAAGTCTTCCCGGATCTTGGACTTCGGTACAGTGTGA	1440
DB	1381	GAAAGGAGACAGGACAGAGACCAAGTCTTCCCGGATCTTGGACTTCGGTACAGTGTGA	1440
QY	1441	AGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCTGTACTTCTGCTCAGTGTCCA	1500
DB	1441	AGCCAGCTCTCCAGGCCACGGAAGTCTCTGAGCAGCTGTACTTCTGCTCAGTGTCCA	1500
QY	1501	GAGTCCATGACAAATTAGTGAATGCGCCAGGCCACTGTGCGCCATGATGAACAGGAGAT	1560
DB	1501	GAGTCCATGACAAATTAGTGAATGCGCCAGGCCACTGTGCGCCATGATGAACAGGAGAT	1560
QY	1561	GAGTGGAGGAGGAGACAGATCACTCGGAAACCTGCTCCAGCGTGAGATGAGCACTCA	1620
DB	1561	GAGTGGAGGAGGAGACAGATCACTCGGAAACCTGCTCCAGCGTGAGATGAGCACTCA	1620
QY	1621	GCCGCGCTCCGCGCAAGAGGTGGAACCTTGAAGAGAGTGGTGAACAGGAGGAGCGG	1680
DB	1621	GCCGCGCTCCGCGCAAGAGGTGGAACCTTGAAGAGAGTGGTGAACAGGAGGAGCGG	1680
QY	1681	CAGGCGATGAAGTCCAGGCGCTGGCCAGCTATCTTTGCTATTGTTGAGGAGATCTAA	1740
DB	1681	CAGGCGATGAAGTCCAGGCGCTGGCCAGCTATCTTTGCTATTGTTGAGGAGATCTAA	1740
QY	1741	CCCCACGTGAGAACCATGTGTGGAGAAATGAGAGGAGAGAGAAATCCACAGTTCCTGA	1800
DB	1741	CCCCACGTGAGAACCATGTGTGGAGAAATGAGAGGAGAGAGAAATCCACAGTTCCTGA	1800
QY	1801	TAGTCTCATTTGAGCTCTCTGATGATCTTTCTGAGCTGTGTTCTCTGAGCTTTT	1860





QY 1795 TCCTGATAGTCTCATTTGAGCTCCCTGGATCCAGTCTTTTCTGAAGCTGTGTTCTCTGG 1854  
Db 1681 TCCTGATAGTCTCATTTGAGCTCCCTGGATCCAGTCTTTTCTGAAGCTGTGTTCTCTGG 1740  
QY 1855 ACTTTTCATGATGTGAGCCATAAATGCTTTTCATTCCTTG 1896  
Db 1741 ACTTTTCATGATGTGAGCCATAAATGCTTTTCATTCCTTG 1782

RESULT 4  
US-10-070-255-6  
; Sequence 6, Application US/10070255  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolay  
; APPLICANT: SINHA, Indranil  
; APPLICANT: LEU, Stefan  
; FILE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE  
; TITLE REFERENCE: WALLACH=28  
; CURRENT APPLICATION NUMBER: US/10/070,255  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/IL00/00517  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: IL 131719  
; PRIOR FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 2873  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-070-255-6

Query Match 81.6%; Score 1555; DB 47; Length 2873;  
Best Local Similarity 99.6%; Pred. No. 0;

Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 118 ATGAGCGGATCNCNGAACNATGACAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGCA 177  
Db 1 ATGAGCGGATCACAGAAACAATGACAAAGACAAATTTCTGCTGGAGCGACTGCTGGATGCA 60  
QY 178 GTGAAACAGTCCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 237  
Db 61 GTGAAACAGTCCAGATCCGCTTTGGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 120  
QY 238 AGGTTCACCTGTCTGTGTGCCAGTTTGAAGCCCTCTGACAGCATGCGTTGAAGAGAGT 297  
Db 121 AGGTTCACCTGTCTGTGTGCCAGTTTGAAGCCCTCTGACAGCATGCGTTGAAGAGAGT 180  
QY 298 CGAGGATTTGGCACTCAGAGCGGACGATCAAGCAGGCGCGGCTTTGCCAGCAAAACC 357  
Db 181 CGAGGATTTGGCACTCAGAGCGGACGATCAAGCAGGCGCGGCTTTGCCAGCAAAACC 240  
QY 358 GAAACAGAGCCGCTGTCTGTGTAAGTCTGAGGAGGTCTCAACAAAGCAGAGCTGCAG 417  
Db 241 GAAACAGAGCCGCTGTCTGTGTAAGTCTGAGGAGGTCTCAACAAAGCAGAGCTGCAG 300  
QY 418 CGCTTCTACTCCCTGGCGCACATCGCTCAGACGTGGCGGGGTGCGGCTGCGCTGCGC 477  
Db 301 CGCTTCTACTCCCTGGCGCACATCGCTCAGACGTGGCGGGGTGCGGCTGCGCTGCGC 360  
QY 478 TGTCCCTCAACGAACACTCCCTGGAGCGGTACTGTCACATGCTCTGCGGCGGCGCTGC 537  
Db 361 TGTCCCTCAACGAACACTCCCTGGAGCGGTACTGTCACATGCTCTGCGGCGGCGCTGC 420  
QY 538 AGGTTCAGCACTTTTATGAAGTCTGTTTGTGATGATGAAGAAAGTCCAGTATG 597  
Db 421 AGGTTCAGCACTTTTATGAAGTCTGTTTGTGATGATGAAGAAAGTCCAGTATG 480  
QY 598 CTTCTACCATGGCAGAGGTCTGAACCTCATCTCTTTGGGATTAATCATCGACAAAG 657  
Db 481 CTTCTACCATGGCAGAGGTCTGAACCTCATCTCTTTGGGATTAATCATCGACAAAG 540

QY 658 GATTTGAACGGGACAGAGTAAGTTTGTCTCCACCCGTTTTCAGACCTCTTAAAGGAGTCAACG 717  
Db 541 GATTTGAACGGGACAGAGTAAGTTTGTCTCCACCCGTTTTCAGACCTCTTAAAGGAGTCAACG 600  
QY 718 CAGAACGTGACCTCTCTTGTGAAGAGTCCACGCAAGAGTGCAGCAGCTGTTTCAGGGAG 777  
Db 601 CAGAACGTGACCTCTCTTGTGAAGAGTCCACGCAAGAGTGCAGCAGCTGTTTCAGGGAG 660  
QY 778 ATCACAGCCTCTCTGCGCTTCCATCTCATCAAACTTGAACAGGAGACCGA-CCCTTG 836  
Db 661 ATCACAGCCTCTCTGCGCTTCCATCTCATCAAACTTGAACAGGAGACCGACCCCTTG 720  
QY 837 CCTGTCTGTCCAGGAATGTCACTGCTGATGAGGAAGTGCAGAAATGCAAAAGAGGAGGAGAA 896  
Db 721 CCTGTCTGTCCAGGAATGTCACTGCTGATGAGGAAGTGCAGAAATGCAAAAGAGGAGGAGAA 780  
QY 897 AAGAAAGTGAACCAATAATCTCATTTGATGATGAGGAAGTGCAGAAATCTCTGGGAC 956  
Db 781 AAGAAAGTGAACCAATAATCTCATTTGATGATGAGGAAGTGCAGAAATCTCTGGGAC 840  
QY 957 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTTCAGAGGACAACTCCGACCGCTCTCT 1016  
Db 841 GTGTTTAAAAAGACACCTGGGGCAGGGGAGAGCTTCAGAGGACAACTCCGACCGCTCTCT 900  
QY 1017 GTCAATATCATGTCCGCTTTTGAAGCCCTTCCGGGCTTAACCTCAATGGAA-TCAGAGC 1075  
Db 901 GTCAATATCATGTCCGCTTTTGAAGCCCTTCCGGGCTTAACCTCAATGGAAAGTCAGAGC 960  
QY 1076 AGCAACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGACTT 1135  
Db 961 AGCAACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGACTT 1020  
QY 1136 GATGTGAAAAAGCATCGATGATGAAGATGTCGATGAAACGAAAGATGACGTGTATGGAAC 1195  
Db 1021 GATGTGAAAAAGCATCGATGATGAAGATGTCGATGAAACGAAAGATGACGTGTATGGAAC 1080  
QY 1196 TCATCAGGAAGGAGCAGAGGGGCCACTCGAGTCCGCGGAGGAGCCACTGGAAGGAAAC 1255  
Db 1081 TCATCAGGAAGGAGCAGAGGGGCCACTCGAGTCCGCGGAGGAGCCACTGGAAGGAAAC 1140  
QY 1256 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGGTGCTGCACAAATGACTCCGAC 1315  
Db 1141 ACCTGCTCTCCAGATGACAGCTGGGCTCCGCTGAAGGTGCTGCACAAATGACTCCGAC 1200  
QY 1316 ATCTCTTCTCTGTCTGAGTGGGCTCTCTA CAGCCACAGCATGCCCCCTCGGAAGC 1375  
Db 1201 ATCTCTTCTCTGTCTGAGTGGGCTCTCTA CAGCCACAGCATGCCCCCTCGGAAGC 1260  
QY 1376 CTGGAGAACGGGACAGGACCCAGAGGCCAGTTCCTCCCGATCCTGGACTTCGGTACAGT 1435  
Db 1261 CTGGAGAACGGGACAGGACCCAGAGGCCAGTTCCTCCCGATCCTGGACTTCGGTACAGT 1320  
QY 1436 GTGGAAGCCAGCTCTCAGGCCACGGAAGTCTCTGAGCAGCTGTTA-CTTCTGCTCA 1494  
Db 1321 GTGGAAGCCAGCTCTCAGGCCACGGAAGTCTCTGAGCAGCTGTTAACCTTCTGCTCA 1380  
QY 1495 GTGCCAGAGTCCATGACAAATTAGTAACTGCGCCAGGCCACTGTGGCCATGATGAACAG 1554  
Db 1381 GTGCCAGAGTCCATGACAAATTAGTAACTGCGCCAGGCCACTGTGGCCATGATGAACAG 1440  
QY 1555 AAGGATGAGTGTGAGGAGGAGACAGATCACTCGGAAACCTGCTCCAGCGTGAAGTGGAG 1614  
Db 1441 AAGGATGAGTGTGAGGAGGAGACAGATCACTCGGAAACCTGCTCCAGCGTGAAGTGGAG 1500  
QY 1615 CACTCAGCCCGCTCCGGCAAGAGGTGGACACCTTGAAGAGAGGTGGTGGTGAACAGAG 1674  
Db 1501 CACTCAGCCCGCTCCGGCAAGAGGTGGACACCTTGAAGAGAGGTGGTGGTGAACAGAG 1560  
QY 1675 GACCGCAGGGCATGAAGTCCAGGGCGCTGGCCAG 1709  
Db 1561 GACCGCAGGGCATGAAGTCCAGGGCGCTGGCCAG 1595

RESULT 5  
US-10-070-255-5  
; Sequence 5, Application US/10070255  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolay  
; APPLICANT: SINHA, Indranil  
; APPLICANT: LEU, Stefan  
; TITLE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE  
; FILE REFERENCE: WALLACH-28  
; CURRENT APPLICATION NUMBER: US/10/070,255  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/IL00/00517  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: IL 131719  
; PRIOR FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 3139  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-070-255-5

Query Match 81.6%; Score 1555; DB 47; Length 3139;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY	118	ATGACGGGATCNGACNATGACAAAGACAAATTTCTGCTGAGCGAGTGGATGCA	177
DB	1	ATGACGGGATCAGACAAATGACAAAGACAAATTTCTGCTGAGCGAGTGGATGCA	60
QY	178	GTAAACAGTCCAGATCCGCTTTGAGGGAGAAAGAGATTCCTCGGATCCGACAGC	237
DB	61	GTAAACAGTCCAGATCCGCTTTGAGGGAGAAAGAGATTCCTCGGATCCGACAGC	120
QY	238	AGGGTCACTGTCTGTGCCCCAGTTTGAAGCCGCTCTGACAGATGGCTTGAAGAGAGT	297
DB	121	AGGGTCACTGTCTGTGCCCCAGTTTGAAGCCGCTCTGACAGATGGCTTGAAGAGAGT	180
QY	298	CGAGGATGGACATCAAGCGGCGAGGATCAAGCAGCGAGCGGGCTTTGCGACAAACC	357
DB	181	CGAGGATGGACATCAAGCGGCGAGGATCAAGCAGCGAGCGGGCTTTGCGACAAACC	240
QY	358	GAACACAGAGCCGTGTTCTGTGTAAGAGAGTCTCAACAGCAGCAGCTGCAG	417
DB	241	GAACACAGAGCCGTGTTCTGTGTAAGAGAGTCTCAACAGCAGCAGCTGCAG	300
QY	418	CGCTTCTACTCTGCGCCACATCGCTCAGACGTGGCGGGGTGCGGCTGGCTGCGC	477
DB	301	CGCTTCTACTCTGCGCCACATCGCTCAGACGTGGCGGGGTGCGGCTGGCTGCGC	360
QY	478	TGTGCGCTCAACGAACCTCCCTGGAGCGGTACTGACATGCTCTGCGCGAGCCGCTGC	537
DB	361	TGTGCGCTCAACGAACCTCCCTGGAGCGGTACTGACATGCTCTGCGCGAGCCGCTGC	420
QY	538	AGGCTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGTCCAGTATG	597
DB	421	AGGCTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGTCCAGTATG	480
QY	598	CTTCTTACCATGTCAGCAGGTCTGAATCCATCTCTTTTGGCATTAACATCGACAAAG	657
DB	481	CTTCTTACCATGTCAGCAGGTCTGAATCCATCTCTTTTGGCATTAACATCGACAAAG	540
QY	658	GATTGAAACGGGAGAGTAAAGTTTGTCCACCCTTTTCAACCTTTAAAGAGTCAACG	717
DB	541	GATTGAAACGGGAGAGTAAAGTTTGTCCACCCTTTTCAACCTTTAAAGAGTCAACG	600
QY	718	CAGAACTGACCTCTTGTGTAAGAGTCCACCAAGAGTGAAGAGCTGTTTCAAGGAG	777
DB	601	CAGAACTGACCTCTTGTGTAAGAGTCCACCAAGAGTGAAGAGCTGTTTCAAGGAG	660

## RESULT 6

US-09-629-469A-14016  
; Sequence 14016, Application US/09629469A  
; GENERAL INFORMATION:  
; APPLICANT: OTA, TOSHIO  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: NISHIKAWA, TETSUO  
; APPLICANT: HAYASHI, KOJI

QY	778	ATCAGAGCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCCTTG	836
DB	661	ATCAGAGCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGACCGA-CCCTTG	720
QY	837	CTGTCTGTCTCAGGAATGTCAAGTGTGATGCCAAATGCAAAAGAGCGGAAGAGAA	896
DB	721	CTGTCTGTCTCAGGAATGTCAAGTGTGATGCCAAATGCAAAAGAGCGGAAGAGAA	780
QY	897	AAGAAAGTGAACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTG	956
DB	781	AAGAAAGTGAACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTG	840
QY	957	GTGTTTAAAAAGACACCTGGGCGAGGAGAGCTCAGAGACAACTCCGACCGCTCTCT	1016
DB	841	GTGTTTAAAAAGACACCTGGGCGAGGAGAGCTCAGAGACAACTCCGACCGCTCTCT	900
QY	1017	GTCAATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAACTCCAAATGGA-TCAGAGC	1075
DB	901	GTCAATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAACTCCAAATGGA-TCAGAGC	960
QY	1076	AGCAACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT	1135
DB	961	AGCAACTCATGGAATAATGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT	1020
QY	1136	GATGGAAGAGATCGATGATGAAGATGGAATGGAATAAAGAGATGACGTGTATGGAAC	1195
DB	1021	GATGGAAGAGATCGATGATGAAGATGGAATGGAATAAAGAGATGACGTGTATGGAAC	1080
QY	1196	TCATCAGGAAGAGACACAGGGGCCACTCGGAGTCCGCCGAGAGCCACTTGGAGGGAAC	1255
DB	1081	TCATCAGGAAGAGACACAGGGGCCACTCGGAGTCCGCCGAGAGCCACTTGGAGGGAAC	1140
QY	1256	ACCTGCTCTCCAGATGACAGTGGGCTCCGCTGAAGGTGTCGACAAATGATCTCGAC	1315
DB	1141	ACCTGCTCTCCAGATGACAGTGGGCTCCGCTGAAGGTGTCGACAAATGATCTCGAC	1200
QY	1316	ATCCTCTTCCCTGTCAAGTGGGCTCTTACAGCCAGAGATGCCCGCTCGGAAGC	1375
DB	1201	ATCCTCTTCCCTGTCAAGTGGGCTCTTACAGCCAGAGATGCCCGCTCGGAAGC	1260
QY	1376	CTGGAGAACGGACAGGACAGAGGACACGTTCTCCCGATCCTCGACTTCGGTACAGT	1435
DB	1261	CTGGAGAACGGACAGGACAGGACACGTTCTCCCGATCCTCGACTTCGGTACAGT	1320
QY	1436	GTGGAGAACGAGCTCTCAGGCCACGGAAGTCTCTGAGCAGCGCTGTTA-CTTCTGCTCA	1494
DB	1321	GTGGAGAACGAGCTCTCAGGCCACGGAAGTCTCTGAGCAGCGCTGTTA-CTTCTGCTCA	1380
QY	1495	GTGCCAGAGTCCATGACAAATTAGTGAATGCGGCCAGGCCACTGTGGCCATGATGAACAGG	1554
DB	1381	GTGCCAGAGTCCATGACAAATTAGTGAATGCGGCCAGGCCACTGTGGCCATGATGAACAGG	1440
QY	1555	AAGGATGAGCTGGAGGAGAGACAGATCACTGCGGAAACCTGCTCAGCGGTGAGATGGAG	1614
DB	1441	AAGGATGAGCTGGAGGAGAGAGACAGATCACTGCGGAAACCTGCTCAGCGGTGAGATGGAG	1500
QY	1615	CACTCAGCGCGCTCCGCGAAGAGGTGGACACTTGTGAAAGAGAGGTGGCTGAACAGGAG	1674
DB	1501	CACTCAGCGCGCTCCGCGAAGAGGTGGACACTTGTGAAAGAGAGGTGGCTGAACAGGAG	1560
QY	1675	GAGCGGAGGAGATGAAGTCCAGCGCTGGCCAG	1709
DB	1561	GAGCGGAGGAGATGAAGTCCAGCGCTGGCCAG	1595







```

; SOFTWARE: pt_CT_genes Version 1.01

```

; SEQ ID NO 9141

; LENGTH: 1516

; TYPE: DNA

ORGANISM: Homo sapiens

; ORGANISM: HOMO SAPIENS  
: FEATURE:

```

; FEATURE:
: NAME/KEY: misc feature

```

```

; NAME/KEY: misc_feature
: LOCATION: (1102) (1576)

```

```

; LOCATION: (1196)...(1516)
; OTHER INFORMATION: similar to gi1710216 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3c00, default parameters

```

US-09-488-725B-9141

Query Match 57.3%; Score 1092.4; DB 22; Length 1516;

Best Local Similarity 97.6%; Pred. No. 7e-259;

3000 Local Similarity 37.0%; Area: NO. 78-255,  
 Matches 1137; Conservative 0; Mismatches 23; Indels 5; Gaps 3;

OV 77 GCCGGCGCGCGCGCGCAGCCCGGCCGGGAGGCGCCATGAGCGGATCNCNGAACN 136

QY / / GCGGCGCGCGCGCANGCACCGGCCGGGGGANAGGCNCCATGAGCGGATCNCNGAACN I36

D<sub>b</sub>

27 GGGGCGGGCGGAGGCACGGCCCGGGGAGAGGCACCATCAGCGCATCACACA 86

DB 27 GGC GGC GCG CCG CAG CAC CGG CCG GAG GAG CAC CAT GAG CGG ATCAC AGAACA 86

[illegible]

QY 13/ ATGACAAAAGACAAATTTCGCTGGAGCGACTGCTGGATGCAGTGAAACAGTGCCAGATCC 196

82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087

D5 87 ATGACAAAAGACAAATTTCCTGCTGGAGCGACTGCTGGATGCAGTGAAACAGTGCCAGATCC 146

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

QY 197 GCTTTNGAGGGAGAAAGGAGATTGCCCTCGGATTCGGACAGCAGGGTCACCTGCTGTGTG 256

100

Db 147 GCTTGGAGGAGAAAGGAGATTGCTCGGATTCCGACAGCAGGGTCACCTGTCTGTGTG 206

QY 257 CCCAGTTTGAAGCCGTCCTGCAGCATGGCTTGAAGAGGAGTCGAGGATTTGGCACTCACAG 316

\_\_\_\_\_

**D<sub>b</sub>** 207 CCCAGTTGAAGCCGTCCTGCAGCATGGCTTTGAAGAGGAGTCGAGGATTGGCACTCACAG 266

[illegible]

QY 317 CGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 376

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

db 267 CGGCAGCGATCAAGCAGGCAGCGGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 326

**THE UNIVERSITY OF CHICAGO**

OV 377 GGTACTACGTGAAGGAGGTCCTCAACAAGCAGAGCTGCAGCGCTTCTACTCCCTGCGCC 436

[illegible]

db 327 GGTACTACGTGAAGGAGGTCTCAACAAGCACGAGCTGCAGCGCTTCTACCTCCCTGGCC 386

527 GGTAACCTACGTGATGGAGGTCCTCAGACAAAGCAGCGTGCAGCGCTTCCTACCTCCCTGGCGCC 386

Ov 437 ACATCGCCTCAGACGCTGGGCCGGGGTCCGGCCCTGGCCCTGTGCCCTCAACGAACACT 496

[illegible]

387 ACATCGCCCTCAGACGCTGGGCGGGGTCGGGGCTGGGCTGTGGCCCTCAACGAACACT 446

22 207 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1



317 CGGAGCGATCAAGCAGCAGCGGCTTTGCCAGCAAAACCGAACAGAGCCCGTGTCT 376  
318 |||||  
319 |||||  
320 |||||  
321 |||||  
322 |||||  
323 |||||  
324 |||||  
325 |||||  
326 CGGAGCGATCAAGCAGCAGCGGCTTTGCCAGCAAAACCGAACAGAGCCCGTGTCT 326  
327 |||||  
328 |||||  
329 |||||  
330 |||||  
331 |||||  
332 |||||  
333 |||||  
334 |||||  
335 |||||  
336 GGTACTACGTGAAGGAGTCTCTCAACAGCAGCGTGCAGCGTCTTACTTCCCTGGCC 436  
337 |||||  
338 |||||  
339 |||||  
340 |||||  
341 |||||  
342 |||||  
343 |||||  
344 |||||  
345 |||||  
346 GGTACTACGTGAAGGAGTCTCTCAACAGCAGCGTGCAGCGTCTTACTTCCCTGGCC 386  
347 |||||  
348 |||||  
349 |||||  
350 |||||  
351 |||||  
352 |||||  
353 |||||  
354 |||||  
355 |||||  
356 ACATCGCTCAGAGTGGGCGGGGTGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 496  
357 |||||  
358 |||||  
359 |||||  
360 |||||  
361 |||||  
362 |||||  
363 |||||  
364 |||||  
365 |||||  
366 ACATCGCTCAGAGTGGGCGGGGTGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 446  
367 |||||  
368 |||||  
369 |||||  
370 |||||  
371 |||||  
372 |||||  
373 |||||  
374 |||||  
375 |||||  
376 CCCTGGAGCGTACCTGCACATGCTCTGCGCGACCGCTGCAGGCTGAGCACTTTTATG 556  
377 |||||  
378 |||||  
379 |||||  
380 |||||  
381 |||||  
382 |||||  
383 |||||  
384 |||||  
385 |||||  
386 CCCTGGAGCGTACCTGCACATGCTCTGCGCGACCGCTGCAGGCTGAGCACTTTTATG 506  
387 |||||  
388 |||||  
389 |||||  
390 |||||  
391 |||||  
392 |||||  
393 |||||  
394 |||||  
395 |||||  
396 AAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGAGCAG 616  
397 |||||  
398 |||||  
399 |||||  
400 |||||  
401 |||||  
402 |||||  
403 |||||  
404 |||||  
405 |||||  
406 AAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGAGCAG 566  
407 |||||  
408 |||||  
409 |||||  
410 |||||  
411 |||||  
412 |||||  
413 |||||  
414 |||||  
415 |||||  
416 GTCTGAATCCATATCTCTTGGATTAACATCGAACAGGATTTGAACGGGAGAGTA 676  
417 |||||  
418 |||||  
419 |||||  
420 |||||  
421 |||||  
422 |||||  
423 |||||  
424 |||||  
425 |||||  
426 GTCCGAATCCATATCTCTTGGATTAACATCGAACAGGATTTGAACGGGAGAGTA 626  
427 |||||  
428 |||||  
429 |||||  
430 |||||  
431 |||||  
432 |||||  
433 |||||  
434 |||||  
435 |||||  
436 AGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGACCTCTTGC 736  
437 |||||  
438 |||||  
439 |||||  
440 |||||  
441 |||||  
442 |||||  
443 |||||  
444 |||||  
445 |||||  
446 AGTTTGTCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGTA---CCTTGC 683  
447 |||||  
448 |||||  
449 |||||  
450 |||||  
451 |||||  
452 |||||  
453 |||||  
454 |||||  
455 |||||  
456 TGAAGGAGTCCACGCAAGAGTGAAGAGTGTTCAGGAGATCACAGCTCTCTTGCCA 796  
457 |||||  
458 |||||  
459 |||||  
460 |||||  
461 |||||  
462 |||||  
463 |||||  
464 |||||  
465 |||||  
466 TGAAGGAGTCCACGCAAGAGTGAAGAGTGTTCAGGAGATCACAGCTCTCTTGCCA 743  
467 |||||  
468 |||||  
469 |||||  
470 |||||  
471 |||||  
472 |||||  
473 |||||  
474 |||||  
475 |||||  
476 TCTCCATCTCTCAATCAACCTGAACAGGAGCGA---CCCTTGCCTGTGTCAGGAGATG 855  
477 |||||  
478 |||||  
479 |||||  
480 |||||  
481 |||||  
482 |||||  
483 |||||  
484 |||||  
485 |||||  
486 TCTCCATCTCTCAATCAACCTGAACAGGAGCGA---CCCTTGCCTGTGTCAGGAGATG 803  
487 |||||  
488 |||||  
489 |||||  
490 |||||  
491 |||||  
492 |||||  
493 |||||  
494 |||||  
495 |||||  
496 TCAGTGTCTGATGCCAAATGCAAAAGAGGAGCGGAAGAAAGAAAGTGAACCAATAA 915  
497 |||||  
498 |||||  
499 |||||  
500 |||||  
501 |||||  
502 |||||  
503 |||||  
504 |||||  
505 |||||  
506 TCAGTGTCTGATGCCAAATGCAAAAGAGGAGCGGAAGAAAGAAAGTGAACCAATAA 863  
507 |||||  
508 |||||  
509 |||||  
510 |||||  
511 |||||  
512 |||||  
513 |||||  
514 |||||  
515 |||||  
516 TCTCATTTGATGATGAGGAAGATGACGAACTCTGGGAGCGTGTGTTAAAGACACCTG 975  
517 |||||  
518 |||||  
519 |||||  
520 |||||  
521 |||||  
522 |||||  
523 |||||  
524 |||||  
525 |||||  
526 ACTCATTTGATGATGAGGAAGATGACGAACTCTGGGAGACATGTTTAAAGACACCTG 923  
527 |||||  
528 |||||  
529 |||||  
530 |||||  
531 |||||  
532 |||||  
533 |||||  
534 |||||  
535 |||||  
536 GGGCAGGGAGAGCTCAGAGGAACTCGACCGCTCTCTGTCAATATCATGTCCGCT 1035  
537 |||||  
538 |||||  
539 |||||  
540 |||||  
541 |||||  
542 |||||  
543 |||||  
544 |||||  
545 |||||  
546 GGGCAGGGAGAGCTCAGAGGAACTCGACCGCTCTCTGTCAATATCATGTCCGCT 983  
547 |||||  
548 |||||  
549 |||||  
550 |||||  
551 |||||  
552 |||||  
553 |||||  
554 |||||  
555 |||||  
556 TTGAAAGCCCTTCGGGCTTAATCCAAATGGAA--TCAGAGCAGCAACTCATGGAAATG 1094  
557 |||||  
558 |||||  
559 |||||  
560 |||||  
561 |||||  
562 |||||  
563 |||||  
564 |||||  
565 |||||  
566 TTGAAAGCCCTTCGGGCTTAATCCAAATGGAA--TCAGAGCAGCAACTCATGGAAATG 1043  
567 |||||  
568 |||||  
569 |||||  
570 |||||  
571 |||||  
572 |||||  
573 |||||  
574 |||||  
575 |||||  
576 ATTCCCTGTCTTGAACGGGAGTTGGTACAGAGCTTGAATGAAAGCATCGATG 1154  
577 |||||  
578 |||||  
579 |||||  
580 |||||  
581 |||||  
582 |||||  
583 |||||  
584 |||||  
585 |||||  
586 ATTCCCTGTCTTGAACGGGAGTTGGTACAGAGCTTGAATGAAAGCATCGATG 1103  
587 |||||  
588 |||||  
589 |||||  
590 |||||  
591 |||||  
592 |||||  
593 |||||  
594 |||||  
595 |||||  
596 ATGAGATGTGTGATGAAACGAGAGTACGTGTATGAAACTCATCAGGAAGAGCACA 1214  
597 |||||  
598 |||||  
599 |||||  
600 |||||  
601 |||||  
602 |||||  
603 |||||  
604 |||||  
605 |||||  
606 ATGAGATGTGTGATGAAACGAGAGTACGTGTATGAAACTCATCAGGAAGAGCACA 1163  
607 |||||  
608 |||||  
609 |||||  
610 |||||  
611 |||||  
612 |||||  
613 |||||  
614 |||||  
615 |||||  
616 GGGGCCACTCGGAGTGGCCCGAGAA 1239  
617 |||||  
618 |||||  
619 |||||  
620 |||||  
621 |||||  
622 |||||  
623 |||||  
624 |||||  
625 |||||  
626 GGGGCCACTCGGAGTGGCCCGAGAA 1188

RESULT 11  
US-09-758-472-1043  
; Sequence 1043, Application US/09758472  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PH001  
; CURRENT APPLICATION NUMBER: US/09/758,472

; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 9632  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1043  
; LENGTH: 1137  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (8)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (121)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (184)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (540)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1100)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1111)  
; OTHER INFORMATION: n equals a,t,g, or c  
; US-09-758-472-1043  
  
Query Match 46.4%; Score 883.6; DB 33; Length 1137;  
Best Local Similarity 98.4%; Pred. No. 2.9e-207;  
Matches 949; Conservative 3; Mismatches 6; Indels 6; Gaps 6;  
  
Qy 364 GAGCCCTGTTCTGGTACTACGTGAGAGGTCCTCAACAGCAGAGCTGCGGCTTC 423  
Db 176 GAGCCCTGTTCTGGTACTACGTGAGAGGTCCTCAACAGCAGAGCTGCGGCTTC 234  
  
Qy 424 TACTCCCTGCGGCACATCGCTCAGAGCTGCGGCTGCGGCTGCGCTGCGCTGCGC 483  
Db 235 TACTCCCTGCGGCACATCGCTCAGAGCTGCGGCTGCGGCTGCGCTGCGCTGCGC 294  
  
Qy 484 CTCAACGAACACATCCCTGGAGCGCTACCTGCACATGCTCTTGCGCGACCGCTGCGGCTG 543  
Db 295 CTCAACGAACACATCCCTGGAGCGCTACCTGCACATGCTCTTGCGCGACCGCTGCGGCTG 354  
  
Qy 544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGTCCAGTATGCTTCCT 603  
Db 355 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGTCCAGTATGCTTCCT 414  
  
Qy 604 ACCATGCGAGAGTCTGAACTCCATCTCTTCCGATTAAACATCGACCAACAGGATTTG 663  
Db 415 ACCATGCGAGAGTCTGAACTCCATCTCTTCCGATTAAACATCGACCAACAGGATTTG 474  
  
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTCCAGCTCTTAAAGAGTCAACCGAGAAC 723  
Db 475 AACGGGAGAGTAAGTTTGTCTCCACCGTTTCCAGCTCTTAAAGAGTCAACCGAGAAC 534  
  
Qy 724 GTGACCTCTTGTGAGAGAGTCCACGAGGAGTGGAGAGTGGTTCAGGAGATCACA 783  
Db 535 GTGACNTCTTGTGAGAGAGTCCACGAGGAGTGGAGAGTGGTTCAGGAGATCACA 594  
  
Qy 784 GCCTCTCTGCGCTCTCCATCTCATCAAACTCAACAGGAGACCGA--CCCTTGCCTGTC 842  
Db 595 GCCTCTCTGCGCTCTCCATCTCATCAAACTCAACAGGAGACCGA--CCCTTGCCTGTC 654  
  
Qy 843 GTGTCCAGGAATGTGAGTGTGATGCCAAATGCAAAAGGAGCGGAGAGAAAGAA 902  
Db 655 GTGTCCAGGAATGTGAGTGTGATGCCAAATGCAAAAGGAGCGGAGAGAAAGAA 714  
  
Qy 903 GTGACCAACAATCTCATTTTGTGATGAGGAAGATGAGCAAACTCTGCGGAGCTGTTT 962



```
Db 715 GTGACCAACATAATCTCTTTGATGATGAGGAAGATGACAGAACTCTGGGACRTGT 774
Qy 963 AAAAAGACACCTCGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 775 AAAAAGACACCTCGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 834
Qy 1023 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAA-TCAGAGCAGCAAC 1081
Db 835 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAAAGTCAGAGCAGCAAC 894
Qy 1082 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGACTTGATGTG 1141
Db 895 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGACTTGATGTG 954
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1201
Db 955 AAAAGCATCGATGATGAAGATGKGAATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1014
Qy 1202 GGAAGGAAGCAGGGGCACTCGGAGTCGCCGAGAGCCACTGGGAAGGAAACACCTGC 1261
Db 1015 GGAAGGAAGCAGGGGCACTCGGAGTCG-CCGAGAAGCCACTGGGAAGGAAACACCTGC 1073
Qy 1262 CTCTCCAGATGCACAGCTGGGCTCCGCTGAGGTGCTGCAATGACTCCGACATCCTC 1321
Db 1074 CT-TTCCAGATGCACAGCTGGGCTCCGNTGAAGG-GCTGNACAATGACTCCGACATCCTC 1131
Qy 1322 TTCC 1325
Db 1132 TTCC 1135
```

## RESULT 12

```
US-10-235-926-1043
; Sequence 1043, Application US/10235926
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH001CIN
; CURRENT APPLICATION NUMBER: US/10/235,926
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: 09/758,472
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9632
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 1043
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (184)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (540)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1100)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1111)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-235-926-1043
Query Match 46.4%; Score 883.6; DB 50; Length 1137;
Best Local Similarity 98.4%; Pred. No. 2.9e-207;
Matches 949; Conservative 3; Mismatches 6; Indels 6; Gaps 6;
Qy 364 GAGCCCGTGTTCGTGTAAGAGAGGTCCTCAACAGCAGCAGCTGCAGCGCTTC 423
Db 176 GAGCCCGTNTTCTGTGTAAGAGAGTCTCAACAGCAGCAGCTGCAG-GCTTC 234
Qy 424 TACTCCCTGGCCACATCGCTCTCAGAGCTGGCGGGTTCGGCTCGCTGTGCTC 483
Db 235 TACTCCCTGGCCACATCGCTCTCAGAGCTGGCGGGTTCGGCTCGCTGTGCTC 294
Qy 484 CTCAACGAACACTCCCTGGAGCGCTACTCGACATGCTCTGGCCGACCGCTGAGGCTG 543
Db 295 CTCAACGAACACTCCCTGGAGCGCTACTCGACATGCTCTGGCCGACCGCTGAGGCTG 354
Qy 544 AGCACTTTTATGAAGACTGCTCTTTTGTGATGATGAAGAAAGTCCAGTATCTTCT 603
Db 355 AGCACTTTTATGAAGACTGCTCTTTTGTGATGATGAAGAAAGTCCAGTATCTTCT 414
Qy 604 ACCATGGCAGCAGCTCTGAACCTCCATCTCTTTGCGATTAAACATCGACAACAGGATTG 663
Db 415 ACCATGGCAGCAGCTCTGAACCTCCATCTCTTTGCGATTAAACATCGACAACAGGATTG 474
Qy 664 AACGGGAGAGTAAGTTTGTCTCCACCGTTTCAGACTCTTTAAAGAGTCAACGCAAC 723
Db 475 AACGGGAGAGTAAGTTTGTCTCCACCGTTTCAGACTCTTTAAAGAGTCAACGCAAC 534
Qy 724 GTGACCTCTCTGTAAGAGTCCACGCAAGAGTGGAGCCTGTTTCAGGAGATCACA 783
Db 535 GTGACNTCTTGTGTAAGAGTCCACGCAAGAGTGGAGCCTGTTTCAGGAGATCACA 594
Qy 784 GCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGAGCCGA-CCCTTCCTGTC 842
Db 595 GCCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGAGCCGACCTTCCTGTC 654
Qy 843 GTGTCCAGGAATGTCTAGTGTGATGCAAAATGCAAAAGAGGCGGAAGAAAAGAAA 902
Db 655 GTGTCCAGGAATGTCTAGTGTGATGCAAAATGCAAAAGAGGCGGAAGAAAAGAAA 714
Qy 903 GTGACCAACATAATCTCTTTGATGATGAGAGATGAGCAGACTCTGGGGAGCTGTT 962
Db 715 GTGACCAACATAATCTCTTTGATGATGAGAGATGAGCAGAACTCTGGGGAGCTGTT 774
Qy 963 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 1022
Db 775 AAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAAT 834
Qy 1023 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAA-TCAGAGCAGCAAC 1081
Db 835 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTTAATCTCAATGGAAAGTCAGAGCAGCAAC 894
Qy 1082 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGACTTGATGTG 1141
Db 895 TCATGGAAAATTGATTCCTCTCTTTGAACGGGGAGTTTGGGTACCAAGACTTGATGTG 954
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1201
Db 955 AAAAGCATCGATGATGAAGATGKGAATGAAACGAAAGATGACGTGTATGGAAAATCATCA 1014
Qy 1202 GGAAGGAAGCAGGGGCACTCGGAGTCGCCGAGAGCCACTGGGAAGGAAACACCTGC 1261
Db 1015 GGAAGGAAGCAGGGGCACTCGGAGTCG-CCGAGAAGCCACTGGGAAGGAAACACCTGC 1073
Qy 1262 CTCTCCAGATGCACAGCTGGGCTCCGCTGAGGTGCTGCAATGACTCCGACATCCTC 1321
Db 1074 CT-TTCCAGATGCACAGCTGGGCTCCGNTGAAGG-GCTGNACAATGACTCCGACATCCTC 1131
```

```
QY 1322 TTCC 1325
Db 1132 TTCC 1135

RESULT 13
US-09-629-469A-18174
; Sequence 18174, Application US/09629469A
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/09/629,469A
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18174
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (364)..(954)
US-09-629-469A-18174

Query Match 41.6%; Score 792.8; DB 28; Length 6045;
Best Local Similarity 96.8%; Pred. No. 1.5e-184;
Matches 852; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

QY 363 AGAGCCCGTGTCTGGTACTAGCTGAAGGAGGTCTCTCAACAGCAGAGCTCC-AGCGGT 421
Db 932 AAGCCCGTGTCTGGTACTAGCTGAAGGAGGTCTCTCAACAGCAGAGCTCCAGCGCT 991
QY 422 TCTACTCTCCGCGCCACATCGCCTCAGACGTGGCGGGGGTCCGCGCTGGCTGGT 481
Db 992 TCTACTCTCCGCGCCACATCGCCTCAGACGTGGCGGGGGTCCGCGCTGGCTGGT 1051
QY 482 CCTCAACCAACACTCCCTGGAGCGCTACTCGACATGCTCTGGCCGACCGCTCGAGC 541
Db 1052 CCTCAACCAACACTCCCTGGAGCGCTACTCGACATGCTCTGGCCGACCGCTCGAGC 1111
QY 542 TGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATCTTC 601
Db 1112 TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATCTTC 1171
QY 602 CTACCATGGAGCAGGTCTGAACCTCCATCTCTTTGCGATTAACTCGACCAAGGATT 661
Db 1172 CTACCATGGAGCAGGTCCGAACCTCCATCTCTTTGCGATTAACTTGACCAAGGATT 1231
```

```
QY 662 TGAACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACGAGA 721
Db 1232 TGAACGGGCGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACGAGA 1291
QY 722 ACGTGACCTCTCTGCTGAAGAGTCCACGCAAGAGTGAAGAGCTCTTTCAGGAGATCA 781
Db 1292 ATGTGA--CCTTGTGAAGAGTCCACGCAAGAGTGAAGAGCTCTTTCAGGAGATCA 1348
QY 782 CAGCTCTCTCTGCGGTCTCCATCTCATCAAACTGAACAGGAGACCGA--CCCTTGCCTG 840
Db 1349 CAGCTCTCTCTGCGCATCTCCATCTCATCAAACTGAACAGGAGACCGA--CCCTTGCCTG 1408
QY 841 TCGTGTCCAGGAATGTCAAGTGTGATGCCAAATGCAAAAAGGAGCGGAAGAAAAAGA 900
Db 1409 TCGTGTCCAGGAATGTCAAGTGTGATGCCAAATGCAAAAAGGAGCGGAAGAAAAAGC 1468
QY 901 AAGTGACCAACATAATCTCAATTTGATGATGAGGAAGATGAGCAGAACTCTGGGACCTGT 960
Db 1469 AAGTGACCAACATTTATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGACATGT 1528
QY 961 TTAATAAGACACCTGGGGGAGGAGCTCAGAGGACAACTCCGACCGCTCTCTCTCA 1020
Db 1529 TTAATAAGACACCTGGGGGAGGAGCTCAGAGGACAACTCCGACCGCTCTCTCTCTCA 1588
QY 1021 ATATCATGTCCGCTTTTGAAGGCCCTTCGGGGCCTTAATCCAATGGAA--TCAGAGCAGCA 1079
Db 1589 ATATCATGTCCGCTTTTGAAGGCCCTTCGGGGCCTTAATCCAATGGAAAGTCAGAGCAGCA 1648
QY 1080 ACTCATGGAATAATGATTCCTCTGTTTGAACGGGGAGTTTGGTACAGAGCTTGATG 1139
Db 1649 ACTCGTGAATAATGATTCCTCTGTTTGAACAGGGAGTTTGGGTACAGAGCTTGATG 1708
QY 1140 TGAAGACATCGATGATGAAGTGTGATGGAAGAGAGATGACGTGTATGGAACATCAT 1199
Db 1709 TGAAGACATCGATGATGAAGTGTGATGGAAGAGAGATGGAAGAGAGATGACGTGTATGGAACATCAT 1768
QY 1200 CAGGAAGGAGCAGCAGGGGCCACTCGGAGTCCGCCGAGAA 1239
Db 1769 CAGGACGGAAGCAGCAGGGGCCACTCAGAGTCCGCCGAGAA 1808

RESULT 14
US-10-917-503-18174
; Sequence 18174, Application US/10917503
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
```

```

; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18174
; LENGTH: 6045
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (364)..(954)
US-10-917-503-18174

```

Query Match	41.6%	Score 792.8	DB 65	Length 6045
Best Local Similarity	96.8%	Pred. No. 1.5e-184		
Matches 852	Conservative 0	Mismatches 22	Indels 6	Gaps 4
QY	363	AGAGCCCGTGTCTTGTTACTACGTGAAGGAGGTCTCTCAACAAGCACAGAGTGC-AGCGCT	421	
DB	932	AAAGCCCGTGTCTTGTTACTACGTGAAGGAGGTCTCTCAACAAGCACAGAGTGC	991	
QY	422	TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCGGGGTTCGGCCCTGGCTGCGCTGTG	481	
DB	992	TCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCGGGCGCCACCTGGCTGCGCTGTG	1051	
QY	482	CCCTCAACGAACACTCCCTGGAGGGCTACCTGCACATGCTCTGGCCGACCGCTGCAGGC	541	
DB	1052	CCCTCAACGAACACTCCCTGGAGGGCTACCTGCACATGCTCTGGCCGACCGCTGCAGGC	1111	
QY	542	TGAGCACTTTTATGAAGACTGGTCTTTTGTGATGAAGAAGTGCAGATATGCTTC	601	
DB	1112	TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGAAGAAGTGCAGATATGCTTC	1171	
QY	602	CTACCATGGCAGAGTCTGAATCCATCTCTTTGCGATTAAACATCGACCAACAAAGGATT	661	
DB	1172	CTACCATGGCAGAGTCCGAACTCCATCTCTTTGCGATTAAACATTTGACCAACAAAGGATT	1231	
QY	662	TGAACGGGAGAGTAAGTTTGCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA	721	
DB	1232	TGAACGGGAGAGTAAGTTTGCTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAGA	1291	
QY	722	ACGTGACCTCTTTGCTGAAGGAGTCCACGCAAGGAGTGACGAGCTGTTTCAGGGAGATCA	781	
DB	1292	ATGTGA---CCTTGCTGAAGAGTCCACGCAAGAGTGACGAGCTGTTTCAGGGAGATCA	1348	
QY	782	CAGCCTCTCTGCGGTCTCCATCCTCATCAAACTGAAACAGGAGACCGA-CCTTGTGCTG	840	
DB	1349	CAGCCTCTCTGCCATCTCCATCCTCATCAAACTGAAACAGGAGACCGACCCCTTGCCCG	1408	
QY	841	TGCTGTCCAGGAATGTCAGTCTGATGCCAAATGCAAAAAGGAGCGGAAGAAAGA	900	
DB	1409	TCGTGTCCAGGAATGTCAGTCTGATGCCAAATGCAAAAAGGAGCGGAAGAAAGAAGC	1468	
QY	901	AAGTGACCAACATACTCTCATTTTGATCAGGAAGATGACGAGAACTCTGGGGACGTGT	960	
DB	1469	AAGTGACCAACATACTCTCATTTTGATGAGGAAGATGACGAGAACTCTGGGGACATGT	1528	
QY	961	TTAAAAGACACTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCTCA	1020	
DB	1529	TTAAAAGACACTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCACTCTCTGTCTCA	1588	
QY	1021	ATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAATCTCAATGGA--TCAGAGAGCA	1079	
DB	1589	ATATCATGTCCGCTTTGAAAGCCCTTTCGGGCGAAATCTCAATGGAAGTTCAGAGCAGCA	1648	
QY	1080	ACTCATGGAATAATGATTCCTCTCTTTGAAACGGGAGTTTGGGTACCAAGACTTGATG	1139	
DB	1649	ACTCGTGGAAATATGATTCCTCTCTTTGAAACGGGAGTTTGGGTACCAAGACTTGATG	1708	
QY	1140	TGAAAAGCATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGGAACCTCAT	1199	

Db	1709	TGAAAAGCATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGGAAACTCAT	1766
Qy	1200	CAGGAGCAAGCACAGGGCCACTCGGAGTTCGCCCCAGAGAA	1239
Db	1769	CAGGACGAAGCAACAGGGCCACTTCAGAGTCGCCCCAGAGAA	1808
<p>RESULT 15</p> <p>US-60-212-356-340</p> <p>; Sequence 340, Application US/60212356</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Beasley, Ellen</p> <p>; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,</p> <p>; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE</p> <p>; TITLE OF INVENTION: PROTEINS, AND USES THEREOF</p> <p>; FILE REFERENCE: CL000677</p> <p>; CURRENT APPLICATION NUMBER: US/60/212,356</p> <p>; CURRENT FILING DATE: 2000-06-19</p> <p>; NUMBER OF SEQ ID NOS: 411</p> <p>; SOFTWARE: FastSeq for Windows Version 4.0</p> <p>; SEQ ID NO 340</p> <p>; LENGTH: 1630</p> <p>; TYPE: DNA</p> <p>; ORGANISM: HUMAN</p> <p>US-60-212-356-340</p>			
<p>Query Match 40.8%; Score 777.4; DB 90; Length 1630;</p> <p>Best Local Similarity 95.3%; Pred. No. 6e-181;</p> <p>Matches 845; Conservative 0; Mismatches 36; Indels 6; Gaps 4;</p>			
Qy	377	GGTACTACGTGAAGGAGGTCTCTCAACAGCACGAGCTGC-AGCGCTTTCTACTCCCTGCGC	435
Db	565	GGTAGTCTCTTCGGGAGGTCTCTCAACAGCACGAGCTGCAAGGCGTCTTACTCCCTGCGC	624
Qy	436	CACATCGCTCAGACGTGGGCGGGGTGCGGCTGCGCTGCGCTCTCAACAGAACAC	495
Db	625	CACATCGCTCAGACGTGGGCGGGGCGCGCTGCGCTGCGCTCTCAACAGAACAC	684
Qy	496	TCCTCGAGCGCTTACCTGCACATGCTCTGGCGGACCGCTGCAGGCTGAGCACTTTTAT	555
Db	685	TCCTCGAGCGCTTACCTGCACATGCTCTGGCGGACCGCTGCAGGCTGAGTACTTTTAT	744
Qy	556	GAAGACTGGTCTTTTGTGATGATGAAGAAGGTCCAGTATGCTTCTTACCATGGCAGCA	615
Db	745	GAAGACTGGTCTTTTGTGATGATGAAGAAGGTCCAGTATGCTTCTTACCATGGCAGCA	804
Qy	616	GGTCTGAATCCCATACTCTTTTGGATTAACATCGACAAACAGGATTTGAAACGGGAGAGT	675
Db	805	GGTCCGAACTCCCATACTCTTTTGGATTAACATTTGACAAACAGGATTTGAAACGGGAGAGT	864
Qy	676	AGGTTTGCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGACCTCTCTTG	735
Db	865	AGGTTTGCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGAACTGAGTGA--CTTTG	921
Qy	736	CTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACAGCCTCTCTTGCC	795
Db	922	CTGAAGGAGTCCACGCAAGGAGTGAGCAGCGTGTTCAGGGAGATCACAGCCTCTCTTGCC	981
Qy	796	GTCTCCATCTCATCAAACTGAAACAGGAGACCGA-CCCTTGCCTGTGTGTGCCAGGAAT	854
Db	982	ATCTCCATCTCATCAAACTGAAACAGGAGACCGACCCCTTGCCTGTGTGTGCCAGGAAT	1041
Qy	855	GTCAAGTCTGATGCCAAATGCAAAAGGAGCGGAAGAGAAAGAAAGTGAACCAACATA	914
Db	1042	GTCAAGTCTGATGCCAAATGCAAAAGGAGCGGAAGAGAAAGAAAGTGAACCAACATA	1101
Qy	915	ATCTCAATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTAAAAAGACACCT	974
Db	1102	ATCTCAATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTAAAAAGACACCT	1161
Qy	975	GGGGCAGGGGAGGCTCAGAGACAACTCCGACCGCTCCTCTGTCAATATCATGTCCGCC	1034
Db	1162	GGGGCAGGGGAGGCTCAGAGACAACTCCGACCGCTCCTCTGTCAATATCATGTCCGCC	1221

Mon Jun 13 13:27:09 2005

Qy	1035	TTTGAAGCCCTTCGGCCCTAACTCCAATGGA-TCAGAGCAGCAACTCATGGAATAATT	1093
Db	1222	TTTGAAGCCCTTCGGCCCTAACTCCAATGGAAGTCAGAGCAGCAACTCGTGGAATAATT	1281
Qy	1094	GATTCCTCTGCTTTGAACGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGCATCGAT	1153
Db	1282	GATTCCTCTGCTTTGAACAGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGCATCGAT	1341
Qy	1154	GATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAGCTCATCAGGAAGGAAGCAC	1213
Db	1342	GATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAAGCTCATCAGGAAGGAAGCAC	1401
Qy	1214	AGGGGCCACTCGGAGTCGCCGAGAGCCACTGGAAGGGGAACACCTG	1260
Db	1402	AGGGGCCACTCAGAGTCGCCGAGAGTGCTGCCATGGCCACGACTG	1448

Search completed: June 11, 2005, 04:00:35  
Job time : 4771.45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 729.11 Seconds  
(without alignments)  
11471.453 Million cell updates/sec

Title: US-09-155-676B-1

Perfect score: 1906

Sequence: 1 cattgggtacgcgggtggcg.....tcattctcttgaaaaaaa 1906

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 12985247 seqs, 2194111492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA New.\*

- 1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq2.\*
- 3: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq.\*
- 8: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq2.\*
- 9: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq3.\*
- 10: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq4.\*
- 11: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq5.\*
- 12: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq2.\*
- 14: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq3.\*
- 15: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1899	99.6	1906	6	US-09-155-676B-1
2	1143	60.0	2248	15	US-60-680-544-46181
3	1143	60.0	2248	15	US-60-680-473-46181
4	791	41.5	876	15	US-60-680-544-46660
5	791	41.5	876	15	US-60-680-473-46660
6	459	24.1	1487	1	PCT-US04-37982-81
7	376	19.7	527690	1	PCT-US04-37982-76
8	333.6	17.5	376	15	US-60-680-544-39116
9	333.6	17.5	376	15	US-60-680-473-39116
10	304.2	16.0	1725	1	PCT-US04-37982-72
11	293.4	15.4	1622	1	PCT-US04-37982-74
12	274.6	14.4	454	8	US-10-450-763-214
13	221	11.6	369941	1	PCT-US04-37982-71
14	206.8	10.8	465	13	US-11-060-756-3234
15	206.8	10.8	465	13	US-11-060-756-7506
16	195.6	10.3	463	8	US-10-450-763-28004
17	135.4	7.1	474	8	US-10-450-763-9353
18	119	6.2	911	8	US-10-450-763-215

19	106.2	5.6	331	12	US-11-124-617-290	Sequence 290, App
20	66.8	3.5	1796	12	US-11-097-143-11288	Sequence 11288, A
21	66.8	3.5	3796	12	US-11-097-143-11287	Sequence 11287, A
22	63.8	3.3	4118	12	US-11-090-997-1431	Sequence 1431, Ap
23	63.4	3.3	3177	15	US-60-680-544-45380	Sequence 45380, A
24	63.4	3.3	3177	15	US-60-680-473-45380	Sequence 45380, A
25	61.6	3.2	1253	6	US-09-155-676B-4	Sequence 4, Appli
26	61.6	3.2	1253	13	US-11-060-756-1553	Sequence 1553, Ap
27	61.6	3.2	1253	13	US-11-060-756-1554	Sequence 1554, Ap
28	61.6	3.2	1253	13	US-11-060-756-5825	Sequence 5825, Ap
29	61.6	3.2	1253	13	US-11-060-756-5826	Sequence 5826, Ap
30	59.4	3.1	5684	9	US-10-472-963-20	Sequence 20, Appl
31	54.8	2.9	36731	11	US-10-940-774A-13770	Sequence 13770, A
32	54.8	2.9	79977	2	PCT-US04-42189-58	Sequence 58, Appl
33	54.6	2.9	4153	12	US-11-090-997-343	Sequence 343, App
34	54.4	2.9	601	11	US-10-940-774A-69138	Sequence 69138, A
35	54.4	2.9	813	8	US-10-450-763-21166	Sequence 21166, A
36	54.4	2.9	188682	12	US-11-112-908-23	Sequence 23, Appl
37	52.4	2.7	537	15	US-60-680-544-41777	Sequence 41777, A
38	52.4	2.7	537	15	US-60-680-473-41777	Sequence 41777, A
39	52.2	2.7	421	12	US-11-000-688A-1575	Sequence 1575, Ap
40	52	2.7	1292	15	US-60-680-544-38736	Sequence 38736, A
41	52	2.7	1292	15	US-60-680-473-38736	Sequence 38736, A
42	52	2.7	1297	15	US-60-680-544-1851	Sequence 1851, Ap
43	52	2.7	1297	15	US-60-680-473-1851	Sequence 1851, Ap
44	50.6	2.6	16420	11	US-10-990-328A-94349	Sequence 94349, A
45	50	2.6	9811	11	US-10-940-774A-12759	Sequence 12759, A

#### ALIGNMENTS

#### RESULT 1

US-09-155-676B-1

Sequence 1 Application US/09155676B

GENERAL INFORMATION:

APPLICANT: WALLACH, David

APPLICANT: MALININ, Nikolai

APPLICANT: BOLDIN, Mark

APPLICANT: KOVALENKO, Andrei

APPLICANT: METT, Igor

TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR

FILE OF INVENTION: PREPARATION AND USE

FILE REFERENCE: WALLACH=21

CURRENT FILING DATE: 1999-01-04

PRIOR FILING DATE: 1999-01-04

PRIOR APPLICATION NUMBER: US/09/155,676B

PRIOR FILING DATE: 1997-04-01

PRIOR APPLICATION NUMBER: IL 117800

PRIOR FILING DATE: 1996-04-02

PRIOR APPLICATION NUMBER: IL 119133

PRIOR FILING DATE: 1996-08-26

NUMBER OF SEQ ID NOS: 22

SOFTWARE: Patent in version 3.3

SEQ ID NO 1

LENGTH: 1906

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (94)..(94)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (110)..(110)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (115)..(115)

OTHER INFORMATION: n is a, c, g, or t

FEATURE:

NAME/KEY: misc feature

LOCATION: (129)..(129)

```

; OTHER INFORMATION: n is a, c, g, or t
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (131)..(131)
; OTHER INFORMATION: n is a, c, g, or t
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (136)..(136)
; OTHER INFORMATION: n is a, c, g, or t
;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (202)..(202)
; OTHER INFORMATION: n is a, c, g, or t
;
; US-09-155-676B-1

```

Query Match	99.68;	Score 1899;	DB 6;	Length 1906;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches 1906;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	CATTGGGTCACGGGTGGCGCGCTCTAGAAATAGTGGATCCCGCGCTGCAGGAATTCG	60		
DB 1	CATTGGGTCACGGGTGGCGCGCTCTAGAAATAGTGGATCCCGCGCTGCAGGAATTCG	60		
QY 61	ATTGAGGGCCACGAAGCGCGCGCGCGCGCANGCACCGGCCCGGGANAGGNCNCATG	120		
DB 61	ATTGAGGGCCACGAAGCGCGCGCGCGCGCANGCACCGGCCCGGGANAGGNCNCATG	120		
QY 121	AGCGGATCNCNGAACNATGACAAAGAGCAATTTCTGTGGAGCACTGCTGGATGCAAGT	180		
DB 121	AGCGGATCNCNGAACNATGACAAAGAGCAATTTCTGTGGAGCACTGCTGGATGCAAGT	180		
QY 181	AAACAGTCCAGATCCGCTTNGAGGGAGAAAGAGATTGCCTCGGATTCGCACAGCAGG	240		
DB 181	AAACAGTCCAGATCCGCTTNGAGGGAGAAAGAGATTGCCTCGGATTCGCACAGCAGG	240		
QY 241	GTCACTCTCTGTGTGCCAGTGTGAAGCCGCTCTGCACAGTGGCTTGAAGAGGAGTCGA	300		
DB 241	GTCACTCTCTGTGTGCCAGTGTGAAGCCGCTCTGCACAGTGGCTTGAAGAGGAGTCGA	300		
QY 301	GGATTGGCACTCACAGCGGCAGCATCAAGCAGGCGCGGCTTTGCCAGCAAAACCGAA	360		
DB 301	GGATTGGCACTCACAGCGGCAGCATCAAGCAGGCGCGGCTTTGCCAGCAAAACCGAA	360		
QY 361	ACAGAGCCCGTGTCTTGTGTACTACGTGAAGGAGTCTCTCAACAGCAGCAGTGCAGCGC	420		
DB 361	ACAGAGCCCGTGTCTTGTGTACTACGTGAAGGAGTCTCTCAACAGCAGCAGTGCAGCGC	420		
QY 421	TTTCTACTCCCTCGGCCACATCGCCTCAGACGTGGCGCGGGTTCGGCCTTGGCTGCCTGT	480		
DB 421	TTTCTACTCCCTCGGCCACATCGCCTCAGACGTGGCGCGGGTTCGGCCTTGGCTGCCTGT	480		
QY 481	GGCCTTCAGGAACACTCCCTTGAGCGCTACCTGCACATGCTCCTCGCCGACCGCTGCAGG	540		
DB 481	GGCCTTCAGGAACACTCCCTTGAGCGCTACCTGCACATGCTCCTCGCCGACCGCTGCAGG	540		
QY 541	CTGAGCACCTTTTATGAAGACTGGTCTTTTGTGAATGGATGAAGAAAGTTCAGTATGCTT	600		
DB 541	CTGAGCACCTTTTATGAAGACTGGTCTTTTGTGAATGGATGAAGAAAGTTCAGTATGCTT	600		
QY 601	CCTACCATGGCAGCAGGTCTGAACCTCCATCTCTTTGCCATTTAAACATCGACAAACAGGAT	660		
DB 601	CCTACCATGGCAGCAGGTCTGAACCTCCATCTCTTTGCCATTTAAACATCGACAAACAGGAT	660		
QY 661	TTGAACGGGCAGAGTAAAGTTTGTCCACCGTTTCAGACCTTTAAAGGAGTCAACGCAG	720		
DB 661	TTGAACGGGCAGAGTAAAGTTTGTCCACCGTTTCAGACCTTTAAAGGAGTCAACGCAG	720		
QY 721	AACGTACCTCTTGTGTGAGGAGTCCACGCAAGGAGTGCAGACGCTGTTTCAGGAGATC	780		
DB 721	AACGTACCTCTTGTGTGAGGAGTCCACGCAAGGAGTGCAGACGCTGTTTCAGGAGATC	780		
QY 781	ACAGCCTCTCTCTCGGCTCTCCATCTCTCATCAAAACCTTGAAACAGGAGACCCCTTGCCTG	840		

Db	781	ACAGCCTCCTCTGCGGTCTCCATCTCTCATCAAACTTGAAACAGGAGACCGACCCCTTCGCTG	840
Qy	841	TCGTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAAGAGAGCGGAAGAAAAAGA	900
Db	841	TCGTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAAGAGAGCGGAAGAAAAAGA	900
Qy	901	AAGTGACCAACATATCTCATTTGATGATGAGGAAGATGACAGAACTCTGGGAGCGTGT	960
Db	901	AAGTGACCAACATATCTCATTTGATGATGAGGAAGATGACAGAACTCTGGGAGCGTGT	960
Qy	961	TTAAAAGACACACTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGTCTCTGTGTCA	1020
Db	961	TTAAAAGACACACTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGTCTCTGTGTCA	1020
Qy	1021	ATATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAAGTCCAAATGGAATCAGAGACGAA	1080
Db	1021	ATATCATGTCCGCTTTGAAAGCCCTTCGGGCGCTTAAGTCCAAATGGAATCAGAGACGAA	1080
Qy	1081	CTCATGTGAAATTTGATTCCTCTGTCTTTGAAACGGGAGTTTGGGTACCAAGACTTGTGATGT	1140
Db	1081	CTCATGTGAAATTTGATTCCTCTGTCTTTGAAACGGGAGTTTGGGTACCAAGACTTGTGATGT	1140
Qy	1141	GAAAGCATCCATGATGAAGATGTGATGAAACGAAGATGACGTGTATGGAACCTCATC	1200
Db	1141	GAAAGCATCCATGATGAAGATGTGATGAAACGAAGATGACGTGTATGGAACCTCATC	1200
Qy	1201	AGGAGGAAGACACAGGGGCCACTCGAGTGCGCCAGAGAACCTTGGAAAGGGAACACCTG	1260
Db	1201	AGGAGGAAGACACAGGGGCCACTCGAGTGCGCCAGAGAACCTTGGAAAGGGAACACCTG	1260
Qy	1261	CCTCTCCAGATGCACAGTGTGGGCTCCGTGAAAGTGTGCACATGACTCCGACATCCT	1320
Db	1261	CCTCTCCAGATGCACAGTGTGGGCTCCGTGAAAGTGTGCACATGACTCCGACATCCT	1320
Qy	1321	CTTCCCTGTCAAGTGGGCTGCTTACAGCCACAGAGATGCCCCCTCGGAAAGCTGGA	1380
Db	1321	CTTCCCTGTCAAGTGGGCTGCTTACAGCCACAGAGATGCCCCCTCGGAAAGCTGGA	1380
Qy	1381	GAAACGGACAGGACACAGGAGACACGTTCTCCGGATCTCGATCTCGGTACAGTGTGGA	1440
Db	1381	GAAACGGACAGGACACAGGAGACACGTTCTCCGGATCTCGATCTCGGTACAGTGTGGA	1440
Qy	1441	AGCAGCTCTCAGGCCACAGGAGTCTCTGACAGCCTGTTACTTCTGCTCTCAGTGCCA	1500
Db	1441	AGCAGCTCTCAGGCCACAGGAGTCTCTGACAGCCTGTTACTTCTGCTCTCAGTGCCA	1500
Qy	1501	GAGTCCATGACAAATTAGTGAACCTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGAT	1560
Db	1501	GAGTCCATGACAAATTAGTGAACCTGCGCCAGGCCACTGTGGCCATGATGAACAGGAAGAT	1560
Qy	1561	GAGCTGGAGGAGGAACAGATCACTCGGAAACCTGCTCGAGGTGAGATGGAGCACTCA	1620
Db	1561	GAGCTGGAGGAGGAACAGATCACTCGGAAACCTGCTCGAGGTGAGATGGAGCACTCA	1620
Qy	1621	GCCGCGCTCCGCAAGAGGTGCAACCTTTGAAAGGAAGGTGGCTGAAACAGGAGAGCGG	1680
Db	1621	GCCGCGCTCCGCAAGAGGTGCAACCTTTGAAAGGAAGGTGGCTGAAACAGGAGAGCGG	1680
Qy	1681	CAGGGCATGAAGGTCCAGGCGCTGGCAGACTATCTTTGCTATTTTGTGAGGAGATCTTAA	1740
Db	1681	CAGGGCATGAAGGTCCAGGCGCTGGCAGACTATCTTTGCTATTTTGTGAGGAGATCTTAA	1740
Qy	1741	CCCCAGTCAGAACCACTGTGTGTGAGAAATGCGGGAGAGAGAAATCCAACTTCTCTGA	1800
Db	1741	CCCCAGTCAGAACCACTGTGTGTGAGAAATGCGGGAGAGAGAAATCCAACTTCTCTGA	1800
Qy	1801	TAGTCTCATTTGAGCTCCTGGATCCAGTCTTCTTCTGAAAGCTGTGTTTCTCTGACCTTTT	1860
Db	1801	TAGTCTCATTTGAGCTCCTGGATCCAGTCTTCTTCTGAAAGCTGTGTTTCTCTGACCTTTT	1860
Qy	1861	CATGTATGTAGCCCAATAATTGCTTTCATTTCTTGAATAAAAAA	1906
Db	1861	CATGTATGTAGCCCAATAATTGCTTTCATTTCTTGAATAAAAAA	1906

RESULT 2  
US-60-680-544-46181  
; Sequence 46181, Application US/60680544  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary  
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,544  
; CURRENT FILING DATE: 2005-05-13  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 46181  
; LENGTH: 2248  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-60-680-544-46181

Query Match 60.0%; Score 1143; DB 15; Length 2248;  
Best Local Similarity 95.3%; Pred. No. 1.2e-259;  
Matches 1195; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

QY 77 GCCGGCGCGCGCGCANGCACCGGCCCGGGAGAGCNCATGAGCGGATCNCNGAACN 136  
DB 22 GCGCGCGCGCGCGCAGGACACCGGCCCGGGAGAGCACCATGAGCGGATCAGAGAAC 81

QY 137 ATGACAAAAGACATTTCTCTGGAGCGACTGCTGGATGAGTGAACAGTGCAGATCC 196  
DB 82 ATGACAAAAGACATTTCTCTGGAGCGACTGCTGGATGAGTGAACAGTGCAGATCC 141

QY 197 GCTTTNAGGAGGAGAAAGGAGATGCTCGGATTCGACAGCAGAGGTCACCTGCTGTGTG 256  
DB 142 GCTTTNAGGAGGAGAAAGGAGATGCTCGGATTCGACAGCAGAGGTCACCTGCTGTGTG 201

QY 257 CCCAGTTTGAAGCGCTCTCAGCATGCTTGAAGAGAGTGCAGGATTCGCACTCACAG 316  
DB 202 CCCAGTTTGAAGCGCTCTCAGCATGCTTGAAGAGAGTGCAGGATTCGCACTCACAG 261

QY 317 CGCAGGAGATCAAGCAGCAGCGGGCTTTCGACGAAACCGAAACGAGCCCGTGTCT 376  
DB 262 CGCAGGAGATCAAGCAGCAGCGGGCTTTCGACGAAACCGAAACGAGCCCGTGTCT 321

QY 377 GGTACTACGTGAAGGAGTCTCTAAACAGCAGCAGTGCAGCGCTTCTACTCTCGCC 436  
DB 322 GGTACTACGTGAAGGAGTCTCTAAACAGCAGCAGTGCAGCGCTTCTACTCTCGCC 381

QY 437 ACATCGCTCAGAGTGGCGGGGTTCGCGCTGGCTGCGCTTGCCTCAAGCAACACT 496  
DB 382 ACATCGCTCAGAGTGGCGGGGTTCGCGCTGGCTGCGCTTGCCTCAAGCAACACT 441

QY 497 CCTGGAGCGGTACCTGCACATGCTCTGGCGACCGCTCGAGCTGAGCACTTTTATG 556  
DB 442 CCTGGAGCGGTACCTGCACATGCTCTGGCGACCGCTCGAGCTGAGCACTTTTATG 501

QY 557 AAGACTGGTCTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTACCATGGCAGCAG 616  
DB 502 AAGACTGGTCTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTACCATGGCAGCAG 561

QY 617 GTCTGAATCCATCTCTTTGCGATTAACATCGACAAAGGATTTGAACGGGCGAGTA 676  
DB 562 GTCTGAATCCATCTCTTTGCGATTAACATCGACAAAGGATTTGAACGGGCGAGTA 621

QY 677 AGTTTGTCTCCACGGTTTCAGACCTCTTTAAAGAGTCAACGCGAGAGCTGACCTCTTC 736

DB 622 AGTTTGTCTCCACCGTTTCAGACCTCTTAAAGAGTCAACGCGAGAGTGCACCTCTTGC 681

QY 737 TGAAGAGTCCACGCAAGGAGTGCAGAGCTGTTCAGGGAGATCACAGCTCTCTCCCG 796

DB 682 TGAAGAGTCCACGCAAGGAGTGCAGAGCTGTTCAGGGAGATCACAGCTCTCTCCCG 741

QY 797 TCTCCATCTCATCAAACTGAAACGAGAGCCGA-CCCTTGCCTGTCTGTCCAGGAATG 855

DB 742 TCTCCATCTCATCAAACTGAAACGAGAGCCGA-CCCTTGCCTGTCTGTCCAGGAATG 801

QY 856 TCAGTGTCTGATGCAAAATGCAAAAGGAGCGGAGAGAAAGAAAGAGTGCACCAATAA 915

DB 802 TCAGTGTCTGATGCAAAATGCAAAAGGAGCGGAGAGAAAGAAAGTGCACCAATAA 861

QY 916 TCTCATTTGATGATGAGGAAGTGCAGAGCAACTCTCTGGGACGTGTTTAAAGACACTG 975

DB 862 TCTCATTTGATGATGAGGAAGTGCAGAGCAACTCTCTGGGACGTGTTTAAAGACACTG 921

QY 976 GGGCAGGGAGAGCTCAGAGGCAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 1035

DB 922 GGGCAGGGAGAGCTCAGAGGCAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 981

QY 1036 TTGAAAGCCCTTTCGGGCTTAACCTCCAATGGA-TCAGAGCAGCAACTCATGGAATG 1094

DB 982 TTGAAAGCCCTTTCGGGCTTAACCTCCAATGGAAGTGCAGAGCAACTCATGGAATG 1041

QY 1095 ATTCCCTGTCTTTGAAACGGGAGTTCGGGTACCGAGCTTGTGTGAAAGCATCGATG 1154

DB 1042 ATTCCCTGTCTTTGAAACGGGAGTTCGGGTACCGAGCTTGTGTGAAAGCATCGATG 1101

QY 1155 ATGAAGATGTGATGAAACGAAAGATGACGTGTATGAAACTCATCAGGAAGGAGCACA 1214

DB 1102 ATGAAGATGTGATGAAACGAAAGATGACGTGTATGAAACTCATCAGGAAGGAGCACA 1161

QY 1215 GGGGCCACTCGGAGTCCCGGAGCCACTGGAAGGCACTGCCTCTCCAGATGC 1274

DB 1162 GGGGCCACTCGGAGTCCCGGAGGAGTGTGTGTAGGTGGAGTCTCACCGTCC 1221

QY 1275 ACAGTGGGCTCGGCTGAAGGTGCTGCACAATGACTCCGACATCTCTTCCCTG 1328

DB 1222 CCAGGCTGGAGTCCGCTGGCGTGTCTCAGTCACTGCAACCTCCACCTCCCG 1275

RESULT 3  
US-60-680-473-46181  
; Sequence 46181, Application US/60680473  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary  
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; CURRENT FILING DATE: 2005-05-13  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 46181  
; LENGTH: 2248  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-60-680-473-46181

Query Match 60.0%; Score 1143; DB 15; Length 2248;  
Best Local Similarity 95.3%; Pred. No. 1.2e-259;  
Matches 1195; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

QY 77 GCCGGCGCGCGCGCANGCACCGGCCCGGGAGAGCNCATGAGCGGATCNCNGAACN 136





Db 475 GTGTCAGGAATGTCAGTCTGATGCCAAATGCAAAAGGAGCGGAAGAAAGCA 534  
Qy 903 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAGCTGTTT 962  
Db 535 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGAGCTGTTT 594  
Qy 963 AAAAAAGACACTGGGGCAGGGGAGAGCTCGAGAGCAAACTCCGACCGCTCCTCTGTCAAT 1022  
Db 595 AAAAAAGACACTGGGGCAGGGGAGAGCTCGAGAGCAAACTCCGACCGCTCCTCTGTCAAT 654  
Qy 1023 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAA--TCAGAGCAGCAAC 1081  
Db 655 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAAAGTCAAGAGCAGCAAC 714  
Qy 1082 TCATGGAATAATGATTCCTCTGTTTGAACGGGAGGTTTGGGTACCAAGAACTGATGTG 1141  
Db 715 TCGTGGAAATGATTCCTCTGTTTGAACAGGAGTTTGGGTACCAAGAACTGATGTG 774  
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGGAACATCATCA 1201  
Db 775 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGGAACATCATCA 834  
Qy 1202 GGAAGGAGCAGCAGGGGCCACTCGAGTCGCCCGAGAG 1240  
Db 835 GGACGGAAGCAGCAGGGGCCACTCAGAGTCGCCCGAGAG 873

RESULT 5  
US-60-680-473-46660  
; Sequence 46660, Application US/60680473  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to  
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; NUMBER OF SEQ ID NOS: 2005-05-13  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 46660  
; LENGTH: 876  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
US-60-680-473-46660

Query Match 41.5%; Score 791; DB 15; Length 876;  
Best Local Similarity 96.8%; Pred. No. 1.3e-176;  
Matches 851; Conservative 0; Mismatches 20; Indels 8; Gaps 4;  
Qy 364 GAGCCCGTGTCTGTTACTAGTGAAGAGGTCTCTCAACAGCAGCAGCTGACGCGCTTC 423  
Db 1 GAGCCCGTGTCTGTTACTAGTGAAGAGGTCTCTCAACAGCAGCAGCT---GCGCTTC 57  
Qy 424 TACTCCCTGCGCCACATCGCTCAGAGCTGGGCGGGGTGCGCTGCGCTGCTGTGCC 483  
Db 58 TACTCCCTGCGCCACATCGCTCAGAGCTGGGCGGGCGCGCTGCGCTGCTGTGCC 117  
Qy 484 CTCAAGCAACACTCCCTGGAGCGCTACCTGCAATGCTCTGCGCCAGCAGCTCAGAGCTG 543  
Db 118 CTCAAGCAACACTCCCTGGAGCGCTACCTGCAATGCTCTGCGCCAGCAGCTCAGAGCTC 177  
Qy 544 AGCAGCTTTTATGAGACTGCTCTTTGTGATCGATGAGAAAGTCCAGATGCTTCCT 603  
Db 178 AGTACTTTTATGAGACTGCTCTTTGTGATGATGATGAGAGAGGTCCAGATGCTTCCT 237

Qy 604 ACCATGTCAGCAGGTCTGAACCTCATATCTCTTTGCGATTAACTCGACAAACAGGATTTG 663  
Db 238 ACCATGTCAGCAGGTCCGAACCTCATATCTCTTTGCGATTAACTCGACAAACAGGATTTG 297  
Qy 664 AACGGCAGAGTAAGTTTGTCTCCACCGTTTCAAGCTCTTAAAGGAGTCAACGCAAGAC 723  
Db 298 AACGGCAGAGTAAGTTTGTCTCCACCGTTTCAAGCTCTTAAAGGAGTCAACGCAAGAC 357  
Qy 724 GTGACCTCTCTTGTGAAGGAGTCCAGCAAGGAGTGAAGAGCTGTTTTCAGGGAGATCACA 783  
Db 358 GTGA---CCTTGTGAAGGAGTCCAGCAAGGAGTGAAGAGCTGTTTTCAGGGAGATCACA 414  
Qy 784 GCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGAGCCGA-CCCTTGGCTGTGTC 842  
Db 415 GCTCTCTGCGCTCTCCATCTCATCAAACTGAAACAGAGAGCCGACCCCTTGGCCGTC 474  
Qy 843 GTCTCAGGAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 902  
Db 475 GTCTCAGGAAATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 534  
Qy 903 GTGACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGGGGAGCTGTTT 962  
Db 535 GTGACCAACATAATCTCATTTGATGATGAGAAAGATGAGCAGAACTCTGGGGAGCTGTTT 594  
Qy 963 AAAAAAGACACTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 1022  
Db 595 AAAAAAGACACTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCCTCTGTCAAT 654  
Qy 1023 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAA--TCAGAGCAGCAAC 1081  
Db 655 ATCATGTCGGCTTTGAAAGCCCTTCGGGCTAACTCCAAATGAAAGTCAAGAGTCAAGAGCAAC 714  
Qy 1082 TCATGGAATAATGATTCCTCTGTTTGAACGGGAGTTCGGGTACCAAGAACTGATGTG 1141  
Db 715 TCGTGGAAATGATTCCTCTGTTTGAACAGGAGTTTGGGTACCAAGAACTGATGTG 774  
Qy 1142 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGGAACATCATCA 1201  
Db 775 AAAAGCATCGATGATGAAGATGTGGATGAAACGAAGATGACGTGTATGGAACATCATCA 834  
Qy 1202 GGAAGGAGCAGCAGGGGCCACTCGAGTCGCCCGAGAG 1240  
Db 835 GGACGGAAGCAGCAGGGGCCACTCAGAGTCGCCCGAGAG 873

RESULT 6  
PCT-US04-37982-81  
; Sequence 81, Application PC/TUS0437982  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER  
; FILE REFERENCE: CHIR0062-500 (23356.0003)  
; CURRENT APPLICATION NUMBER: PCT/US04/37982  
; PRIOR FILING DATE: 2004-10-21  
; PRIOR FILING DATE: 2003-10-22  
; PRIOR FILING DATE: 2003-10-22  
; PRIOR FILING DATE: 2003-10-21  
; NUMBER OF SEQ ID NOS: 3504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 81  
; LENGTH: 1487  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
PCT-US04-37982-81

Query Match 24.1%; Score 459; DB 1; Length 1487;  
Best Local Similarity 99.8%; Pred. No. 3.4e-98;  
Matches 470; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1240 GCCACTGGAAGGAAACACCTGCTCTCCAGATGCAAGCTGGCTCCGCTGAAGGTGCT 1299

Db	1	GC	CA	CT	TG	GA	AG	GA	CA	CA	CT	TG	CC	CA	GT	CA	CA	GT	GG	CT	CG	CT	GA	AG	GT	GC	60	
Qy	1300	GC	CA	AT	GA	CT	CC	GA	CA	TC	CT	TT	CC	CT	GT	CA	GT	GG	CT	GG	CT	CT	CA	CA	GC	CA	1359	
Db	61	GC	CA	AT	GA	CT	CC	GA	CA	TC	CT	TT	CC	CT	GT	CA	GT	GG	CT	GG	CT	CT	CA	CA	GC	CA	120	
Qy	1360	TG	CC	CC	CT	CG	GA	AG	CC	TG	GA	GA	AA	CG	GA	CA	GA	CA	CA	GT	TT	CT	CC	CG	CA	TC	1419	
Db	121	TG	CC	CC	CT	CG	GA	AG	CC	TG	GA	GA	AA	CG	GA	CA	GA	CA	CA	GT	TT	CT	CC	CG	CA	TC	180	
Qy	1420	TG	GA	CT	TC	GG	TA	CA	GT	GT	GA	AG	CC	AG	CT	CT	CA	GG	CC	AC	GG	AA	GT	CT	CT	GA	CA	1479
Db	181	TG	GA	CT	TC	GG	TA	CA	GT	GT	GA	AG	CC	AG	CT	CT	CA	GG	CC	AC	GG	AA	GT	CT	CT	GA	CA	240
Qy	1480	GT	TA	-	CT	TC	GC	CT	CA	GT	GC	CA	GT	CC	AT	GA	CA	AA	TT	AG	TT	GA	CT	GC	GC	CA	GC	1538
Db	241	GT	TA	AC	CT	TC	GC	CT	CA	GT	GC	CA	GT	CC	AT	GA	CA	AA	TT	AG	TT	GA	CT	GC	GC	CA	GC	300
Qy	1539	GG	CA	T	GA	T	GA	AA	CA	GG	AA	GA	GT	AG	CT	TG	GA	GG	AG	GA	CA	GA	T	CA	CT	GC	GA	1598
Db	301	GG	CA	T	GA	T	GA	AA	CA	GG	AA	GA	GT	AG	CT	TG	GA	GG	AG	GA	CA	GA	T	CA	CT	GC	GA	360
Qy	1599	CG	AC	GT	CA	GA	T	GA	GG	AG	CA	CT	CAG	CC	CG	CT	CC	GG	CA	AG	GT	GG	AC	CA	CT	TG	GA	1658
Db	361	CG	AC	GT	CA	GA	T	GA	GG	AG	CA	CT	CAG	CC	CG	CT	CC	GG	CA	AG	GT	GG	AC	CA	CT	TG	GA	420
Qy	1659	GG	T	GG	CT	GA	CA	GG	AG	AG	CG	CA	GG	CA	TA	GA	AG	GT	TC	AG	GG	CT	GG	CA	GC	AG	1709	
Db	421	GT	GG	CT	GA	CA	GG	AG	AG	CG	CA	GG	CA	TA	GA	AG	GT	TC	AG	GG	CT	GG	CA	GC	AG	471		

```

RESULT 7
; Sequence 76, Application PC/TUS0437982
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER
; FILE REFERENCE: CHIR0062-500 (23356.0003)
; CURRENT APPLICATION NUMBER: PCT/US04/37982
; CURRENT FILING DATE: 2004-10-21
; PRIOR APPLICATION NUMBER: US 10/692,382
; PRIOR FILING DATE: 2003-10-22
; PRIOR APPLICATION NUMBER: US 10/691,209
; PRIOR FILING DATE: 2003-10-21
; NUMBER OF SEQ ID NOS: 3504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 527690
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(527690)
; OTHER INFORMATION: n = A, C, G or T/U
PCT-US04-37982-76

```

	Query Match	19.7%	Score 376;	DB 1;	Length 527690;
	Best Local Similarity	89.1%	Pred. No. 4.8e-78;		
	Matches 417;	Conservative	0;	Mismatches 50;	Indels 1; Gaps 1;
QY	862	CTGATGCCAAATGCAAAAAGGACGGAGAGAAAGAAAGTGCACCAACATATCTCAT	921		
DB	317	CAGATGCCAAATGCAAAAAGGACGGAGAGAAAGAAAGTGCACCAACATATCTCAT	376		
QY	922	TTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTAAAAAGACACCTGGGGCAG	981		
DB	377	TTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTTAAAAAGACACCTGGGGCAG	436		
QY	982	GGGAGAGCTCAGAGACAACTCCGACCGCTCTCTGTCTCAATATCATGTTCGCCCTTTGAAA	1041		
DB	437	GGGAGAGCTCAGAGACAACTCCGACCGCTCTCTGTCTCAATATCATGTTCGCCCTTTGAAA	496		

Qy	1042	GC	CCCTTCGGG	CC	TAACTCCA	TGGAA--TCAG	CAGCAGCA	AACTCATG	AGAAAT	TGAT	TCCC	11001
Db	497	GC	CCCTTCGGG	CC	TAACTCCA	TGGAA	GTGAGCAG	CAACTCATG	AGAAAT	TGAT	TCCC	556
Qy	1101	TG	TCCTTTGA	ACGGGG	AGTTTGG	GTACCA	GAACTTCG	ATGTGAA	AACCAT	CGATGAT	GAA	1160
Db	557	TG	TCCTTTGA	ACGGGG	AGTTTGG	GTACCA	GAACTTCG	ATGTGAA	AACCAT	CGATGAT	GAA	616
Qy	1161	AT	TGGATGAAA	ACGA	AGATGAC	GTATGGA	AACTTCAT	CAGGAAG	GAAGCA	CAGGGCC	1220	
Db	617	AT	TGGATGAAA	ACGA	AGATGAC	GTATGGA	AACTTCAT	CAGGAAG	GAAGCA	CAGGGCC	676	
Qy	1221	ACT	CGAGTTCG	CCGGAAG	CCACTG	GAAGGAA	CACTTGC	CTCTCC	CAGATC	CACAGCT	1280	
Db	677	ACT	CGAGTTCG	CCGGAAG	ATAAGTTT	GTGTGA	AGGTGG	AGTCTAC	CGTCCC	CCAGGC	736	
Qy	1281	GGG	CTCCGT	TGAAG	TGCTG	CAAAAT	GACTCCG	ACATCTCT	TTCCTG	1328		
Db	737	TG	AGTCTCC	TGCTG	ATCTC	AGCTC	ACTGCA	ACCTCC	ACCTCC	CGG	784	

## RESULT 8

```

RES001 8
US-60-544-39116
; Sequence 39116, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; Cynomolgus Monkey Genes and the Use Thereof
; FILE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 39116
; LENGTH: 376
; TYPE: DNA
; ORGANISM: Macaca Mulatta
US-60-680-544-39116

```

Query Match	17.5%;	Score 333.6;	DB 15;	Length 376;
Best Local Similarity	94.7%;	Pred. No. 1e-68;		
Matches 356;	Conservative 0;	Mismatches 19;	Indels 1;	Gaps 1;
QY	864	GATGCCAAATCCAAAAGGAGCGGAAGACGAAAAGAAAGAGTGCACCAACATAATCTCATTT	923	
DB	1	GATGCCAAATGCCAAAAGGAGCGGAAGAGAAAAGAAAGAGTGCACCAATATATCTCATTT	60	
QY	924	GATGATGAGGAAGATGAGCAAACTCTGGGACGCTGTTTAAAAAGACACCTCTGGGCGAGGG	983	
DB	61	GATGATGAGGAGGATGAGCAAACTCTGGGACATTTTAAAAAGACACCTCTGGGCGAGGG	120	
QY	984	GAGAGCTCAGAGGACAACTCCGACCGCTCTCTCTCAATATCATGTCCGCTTTGAAAGC	1043	
DB	121	GAGAGCTCGGAGGACAACTCCGACCGCTCTCTCTGTAATATAATATCAGGCTTTGAAAGC	180	
QY	1044	CCCTTCGGGCTTAAGCTCAATGGAA-TCAGAGCAGCAACTCATCGAAAAATGATTCCTCG	1102	
DB	181	CCCTTTGGGCAAAATTCCAATGGAAAGTCAGAGCAGCAACTCGTCGAAAAATGATTCCTCG	240	
QY	1103	TCCTTTGAACGGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGATCGATGATGAAGAT	1162	
DB	241	TCCTTTGAATGGGAGTTTGGGTACCAGAAGCTTGATGTGAAAAGATCGATGATGAAGAT	300	
QY	1163	GTGGATGAAAACGAAGATGACGTGTATGAAAACCTCATCAGAAAGGAAGCACAGGGGCCAC	1222	
DB	301	GTGGATGAAAACGAAGATGACGTGTATGAAAACCTCATCAGAAAGGAAGCACAGGGGCCAC	360	

QY 1223 TCGAGTCCGCCGAGA 1238  
| | | | | | | | | |  
Db 361 GCAGAGTCACTGAGA 376

## RESULT 9

US-60-680-473-39116  
; Sequence 39116, Application US/60680473  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to  
; FILE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 39116  
; LENGTH: 376  
; TYPE: DNA  
; ORGANISM: Macaca Mulatta  
US-60-680-473-39116

Query Match 17.5%; Score 333.6; DB 15; Length 376;  
Best Local Similarity 94.7%; Pred. No. 1e-68;  
Matches 356; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 864 GATGCCAAATGCAAAAGGAGCGGAGAGCAAGAAAAAGAAAGTACCAACATATCTCATTT 923  
| | | | | | | | | |  
Db 1 GATGCCAAATGCAAAAGGAGCGGAGAGCAAGAAAAAGAAAGTACCAACATATCTCATTT 60  
| | | | | | | | | |  
QY 924 GATGATGAGGAAGATGAGCAGAACTCTGGGACGCTGTTTAAAAAGACACCTGGGGCAGGG 983  
| | | | | | | | | |  
Db 61 GATGATGAGGAAGATGAGCAGAACTCTGGGACATTTTAAAAAGACACCTGGGGCAGGG 120  
| | | | | | | | | |  
QY 984 GAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCATATATATGTCGCGCTTTGAAAGC 1043  
| | | | | | | | | |  
Db 121 GAGAGCTCGAGGACAACTCCGACCGCTCTCTGTCATATATATATGTCGCGCTTTGAAAGC 180  
| | | | | | | | | |  
QY 1044 CCCTTGGGCTTAACCTCAATGGAA-TCAGAGCAGCAACTCATGGAATAATGATTCCTG 1102  
| | | | | | | | | |  
Db 181 CCCTTGGGCTTAATCCAAATGGAAATGCAAGTCAAGCAGCAACTCGTGGAAATGATTCCTG 240  
| | | | | | | | | |  
QY 1103 TCTTTGAACGGGAGTTTGGGTACCAAGCTTGATGTAAGAAAGCATCGATGATGAAGAT 1162  
| | | | | | | | | |  
Db 241 TCTTTGAATGGGAGTTTGGGTACCAAGCTTGATGTAAGAAAGCATCGATGATGAAGAT 300  
| | | | | | | | | |  
QY 1163 GTGGATGAAACGAAGATGACGTGTATGGAACCTCATCAGGAAGGAGCAAGCGGCCAC 1222  
| | | | | | | | | |  
Db 301 GTGGATGAAACGAAGATGACGTGTATGGAACCTCATCAGGAAGGAGCAAGCGGCCAC 360  
| | | | | | | | | |  
QY 1223 TCGAGTCCGCCGAGA 1238  
| | | | | | | | | |  
Db 361 GCAGAGTCACTGAGA 376

## RESULT 10

PCT-US04-37982-74  
; Sequence 74, Application PC/TUS0437982  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER  
; FILE REFERENCE: CHIR0062-500 (23356.0003)  
; CURRENT APPLICATION NUMBER: PCT/US04/37982

; CURRENT FILING DATE: 2004-10-21  
; PRIOR APPLICATION NUMBER: US 10/692,382  
; PRIOR FILING DATE: 2003-10-22  
; PRIOR APPLICATION NUMBER: US 10/691,209  
; PRIOR FILING DATE: 2003-10-21  
; NUMBER OF SEQ ID NOS: 3504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 74  
; LENGTH: 1725  
; TYPE: DNA  
; ORGANISM: Mus musculus  
PCT-US04-37982-74

Query Match 16.0%; Score 304.2; DB 1; Length 1725;  
Best Local Similarity 80.1%; Pred. No. 1.2e-61;  
Matches 387; Conservative 0; Mismatches 83; Indels 13; Gaps 2;

QY 1240 GCACACTGGAGGGAACACCTGCTCCAGATGACACAGCTGGGCTCCGCTGAAGTGCT 1299  
| | | | | | | | | |  
Db 241 GACACTGGAGTGGGAACGCTGCTGCCAGGTCCATGGCTGGGCTCCACTGCAGGTGCT 300  
| | | | | | | | | |  
QY 1300 GCACAA-----TGACTCCGACATCTCTTCCCTGTGAGTGGGCTGCTCTTA 1347  
| | | | | | | | | |  
Db 301 TCATGGAGATCGGATGCTGACACTGATGTACTTCTCCCTGTGAGTGGGCTGCTTA 360  
| | | | | | | | | |  
QY 1348 CAGCCCGAGCAGATGCCCCCTCGAAGCCTGGAGAACGGGACAGGACAGGACCAAGT 1407  
| | | | | | | | | |  
Db 361 TGGTGACGAGATGCCCCCTGTGGAAAGTCTGGAGAAATGGAACAGGAACAGAAACCATAT 420  
| | | | | | | | | |  
QY 1408 TCTCCCGGATCTGGACCTTGGTACAGTGTGGAGGACAGCTCTCCAGGCCACGGAGTCC 1467  
| | | | | | | | | |  
Db 421 AATCCCGAGAGCTGGGCTTTCGGTACAGTAGGGAAGCCAGTTCTCCAGGCCAAGGAAGTCC 480  
| | | | | | | | | |  
QY 1468 TCTGACGAGCTGTGA-CTTCTGCCCTCAGTGCCAGAGTCCATGACAAATGATGAAGTCCG 1526  
| | | | | | | | | |  
Db 481 TCTGACGAGCTGTGTAACCTTCTGCTCAGTGCCTGAGTCCATGACAGTCCATGAGTCCG 540  
| | | | | | | | | |  
QY 1527 CCAGGCCACTGTGGCCATGATGAACAGGAAGGATGAGTGGAGGAGGAGAACAGATCACT 1586  
| | | | | | | | | |  
Db 541 CCAAGCCATCTGTGGCCATGATGAACAGGAAGGATGAATGGAAGAGAGAGAAATGGTCACT 600  
| | | | | | | | | |  
QY 1587 GCGAAACCTGCTCGAGGCTGAGATGAGCACTCAGCGGCTCCGCGCAGAGGTGGACAC 1646  
| | | | | | | | | |  
Db 601 GCGAAATCTCTTTCAGCGTGAGATGGAGCACTCAGCAGCGCTCCGCGCAGAGGTGGATGC 660  
| | | | | | | | | |  
QY 1647 CTTGAAAAAGAAAGTGGCTGAAACAGGAGGAGCGGAGGATCAAGGTCCAGGCGCTGGC 1706  
| | | | | | | | | |  
Db 661 CTTGAGGCGGAAGGTGACAGAACAGGAGGAGGCCCATGTCACAAAGGTCCAGGCACTGGC 720  
| | | | | | | | | |  
QY 1707 CAG 1709  
| | |  
Db 721 CAG 723

## RESULT 11

PCT-US04-37982-72  
; Sequence 72, Application PC/TUS0437982  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc S. Malandro  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS IN CANCER  
; FILE REFERENCE: CHIR0062-500 (23356.0003)  
; CURRENT APPLICATION NUMBER: PCT/US04/37982  
; CURRENT FILING DATE: 2004-10-21  
; PRIOR APPLICATION NUMBER: US 10/692,382  
; PRIOR FILING DATE: 2003-10-22  
; PRIOR APPLICATION NUMBER: US 10/691,209  
; PRIOR FILING DATE: 2003-10-21  
; NUMBER OF SEQ ID NOS: 3504  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 72  
; LENGTH: 1622  
; TYPE: DNA





**THIS PAGE BLANK (USPTO)**

Result No.	Query %			DB	ID	Description
	Match	Score	Length			
1	3058	98.9	604	2	AaW42400	AaW42400 Partial s
2	1594.5	51.6	541	4	AaE00683	AaE00683 Human TNF
3	1537.5	49.7	784	4	AaE00685	AaE00685 Human TNF
4	1537.5	49.7	813	4	AaS00684	AaE00684 Human TNF
5	1196	38.7	375	4	AaB93362	AaB93362 Human pro
6	388	12.5	91	4	ABG00223	ABG00223 Novel hum
7	302	9.8	520	4	ABSB61499	ABSB61499 Drosophil
8	236	7.6	131	4	ABG28013	ABG28013 Novel hum
9	183.5	5.9	720	4	ABSB62508	ABSB62508 Drosophil
10	183	5.9	1019	4	ABSB11328	ABSB11328 Human BG3
11	175.5	5.7	473	6	ADAS4368	ADAS4368 Human pro
12	175.5	5.7	473	7	ADJ70685	ADJ70685 Human hea
13	171	5.5	1006	4	ABG21178	ABG21178 Novel hum
14	160	5.2	2845	8	ADo08046	ADo08046 Mouse pol
15	154	5.0	388	4	ABG21174	ABG21174 Novel hum
16	151	4.9	658	7	ADSo7881	ADSo7881 Novel pro
17	151	4.9	704	2	AaY29517	AaY29517 Human lun
18	151	4.9	704	3	AaB44467	AaB44467 Human lun
19	151	4.9	704	4	AaE13809	AaE13809 Human lun
20	151	4.9	704	7	ADd66499	ADd66499 Human lun
21	151	4.9	704	7	ADSB87753	ADSB87753 Human lun
22	149	4.8	641	8	ADr08868	ADr08868 Human pro
23	145	4.7	546	5	ABSB90404	ABSB90404 Human pol
24	144.5	4.7	342	7	ADd71112	ADd71112 Human int
25	144	4.7	600	2	AaY29861	AaY29861 Human sec

FT /note= "not specified"

FT Misc-difference 549 /note= "not specified"

FT Misc-difference 569 /note= "not specified"

FT Misc-difference 570 /note= "not specified"

XX WO9737016-A1.

XX 09-OCT-1997.

XX 01-APR-1997; 97WO-IL000117.

XX 02-APR-1996; 96IL-00117800.

XX 26-AUG-1996; 96IL-00119133.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX WPI: 1997-503101/46.

XX N-PSDB; AAV03323.

XX DNA encoding tumour necrosis factor receptor-associated factor binding molecule - used for modulation or mediation in cells of the activity of NF-kB.

XX Disclosure; Fig 3D; 127pp; English.

XX The present sequence represents the sequence encoded by the 5' end of clone 9, and is a TRAF2 binding protein. Clone 9 is a partial clone, which lacks most of its 5' end of the coding DNA sequence. A cDNA library prepared from B-cells was screened for proteins that associate with TRAF2, and clone 9 DNA isolated. The clone 9 protein is capable of binding to at least amino acids 222-501 of TRAF2. The TRAF-2 binding proteins can be used for modulation or mediation in cells of the activity of NF-kappaB or any other intracellular signalling activity modulated or mediated by TRAF2. TRAF-binding proteins are especially used for prevention or treatment of pathological conditions associated with NF-kB induction, e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the beta langerhans cells or the pancreas that results in diabetes, the death of cells in graft rejection, the death of oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited T cell suicide which causes proliferation of the AIDS virus and hence the AIDS disease. The proteins are also useful for screening of ligands capable of binding to a protein, which are useful for modulating cellular activity modulated/mediated by TRAF2

XX Sequence 604 AA;

Query Match 98.9%; Score 3058; DB 2; Length 604;

Best Local Similarity 100.0%; Pred. No. 7e-272;

Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TGPXGXWMSGXNKKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVL 61

DB 2 TGPXGXWMSGXNKKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVL 61

QY 62 QHGLKRSRGLATAAIAKQAGFASKTETEPVFWYVYVKEVNLKHELORFYSLRHTASDVG 121

DB 62 QHGLKRSRGLATAAIAKQAGFASKTETEPVFWYVYVKEVNLKHELORFYSLRHTASDVG 121

QY 122 RGRWLRCALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMSLPTMAAGLSILF 181

DB 122 RGRWLRCALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMSLPTMAAGLSILF 181

QY 182 AINIDNKLNGQSKFAFTVSDLLKSTQNTVTSLLKSTQGVSSLPFEITASSAVSILIKP 241

DB 182 AINIDNKLNGQSKFAFTVSDLLKSTQNTVTSLLKSTQGVSSLPFEITASSAVSILIKP 241

QY 242 EQETDPCLCSPGMSVLMFNKRSRRRKKYPTXSHLMNRKMSRTLTGTLKRHLGQGAQR 301

DB 242 EQETDPCLCSPGMSVLMFNKRSRRRKKYPTXSHLMNRKMSRTLTGTLKRHLGQGAQR 301

QY 302 TTPTAPLSICPPLKAPSGLTPMSEQOLMENXFPVFERGVWVPEAXCEKHXRCGXKR 361

DB 302 TTPTAPLSICPPLKAPSGLTPMSEQOLMENXFPVFERGVWVPEAXCEKHXRCGXKR 361

QY 362 RXRVWKLIRKEAQGPGLVAREATGREHLPLDPAQIGSAEAGAAQXLRHPLPCQWRGLLOPS 421

DB 362 RXRVWKLIRKEAQGPGLVAREATGREHLPLDPAQIGSAEAGAAQXLRHPLPCQWRGLLOPS 421

QY 422 RCPPRKPGERDTRGPRSPGWSVQCGSOLSRPKSSEQPVTSASVPESMTISELROAT 481

DB 422 RCPPRKPGERDTRGPRSPGWSVQCGSOLSRPKSSEQPVTSASVPESMTISELROAT 481

QY 482 VAMNRKDELEENSLRNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKVQALASYL 541

DB 482 VAMNRKDELEENSLRNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKVQALASYL 541

QY 542 CYFVRRFXPHVTRMWRNGREKSNSSXXSHLSWISQSLKCLFLWTFHVCEPINCFHSL 601

DB 542 CYFVRRFXPHVTRMWRNGREKSNSSXXSHLSWISQSLKCLFLWTFHVCEPINCFHSL 601

QY 602 KKK 604

DB 602 KKK 604

RESULT 2

AAE00683

ID AAE00683 standard; protein; 541 AA.

XX AC AAE00683;

XX DT 04-JUL-2001 (first entry)

XX DE Human TNF receptor-associated factor (TRAF2) binding protein, IREN.

XX KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor; TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytostatic; tumour; AIDS;

XX KW acquired immune deficiency syndrome; rheumatic disease; apoptosis; autoimmune disease; septic shock; graft-vs-host reaction; inflammation; anorexia; anti-HIV; therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Binding-site 198..388

FT Region 326..331

FT /note= "This motif is also present within 51 amino acid domain spanning amino acids 769-820 of NIK (Nuclear factor (NF)-kappaB inducing kinase) which is essential for IKK-1 (IkappaB Kinase) binding to NIK in a 2-hybrid assay and NF-kappaB activation by NIK over expression"

XX WO200116314-A1.

XX PD 08-MAR-2001.

XX PF 31-AUG-2000; 2000WO-IL0000517.

XX PR 02-SEP-1999; 99IL-00131719.

XX PA (YEDA ) YEDA RES & DEV CO LTD.

XX PI Wallach D, Malinin N, Sinha I, Leu S;

XX WPI: 2001-281387/29.

XX N-PSDB; AAD04333.

XX New DNA sequence encoding Tumor Necrosis Factor receptor associated factor (TRAF) binding proteins (IREN) for treatment or prevention of pathological conditions associated with NF-kappaB induction.



XX Claim 17; Fig 6; 118pp; English.

XX The present sequence is human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein, IREN. A fragment of this IREN molecule is capable of binding to TRAF2 protein at position 225-501. The invention relates to human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their corresponding cDNA molecules. IREN is useful for modulating/mediating the activity of intracellular signalling factor NF (Nuclear Factor)-kappaB or any other the prevention and treatment of a pathological condition associated with NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia, septic shock and graft-vs-host reactions. IREN also plays an important role in the control of inflammation and other non-apoptotic effects of TNF as well as in the control of apoptosis. The invention also relates to method for screening, identifying and producing a molecule capable of modulating activities mediated by IREN. IREN antibodies are useful for the purification of new proteins from different sources, including cell extracts or transformed cell lines, in addition IREN can be used in diagnostic purposes for identifying disorders related to abnormal functioning of cellular effects mediated directly by TRAF proteins

XX Sequence 541 AA;

Query Match 51.6%; Score 1594.5; DB 4; Length 541;  
Best Local Similarity 61.8%; Pred. No. 2.5e-137;  
Matches 366; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

Qy 9 MSGSXNDKROFLERLLDAVKQCIQIRFGKKEIASDSRVTCICAQFEAVLQGLKRS 68  
Db 1 MSGSQNDKROFLERLLDAVKQCIQIRFGKKEIASDSRVTCICAQFEAVLQGLKRS 60

Qy 69 RGLALTAATAIKQAAGPASKTETEPFVYVYKEVNLKHELQRFYSLRHIASDVGRGAWLR 128  
Db 61 RGLALTAATAIKQAAGPASKTETEPFVYVYKEVNLKHELQRFYSLRHIASDVGRGAWLR 120

Qy 129 CALNEHSLERYLHMLADRCLSTFYEDWSFVMDERSSMLPTWAAGNSILFPAINDNK 188  
Db 121 CALNEHSLERYLHMLADRCLSTFYEDWSFVMDERSSMLPTWAAGNSILFPAINDNK 180

Qy 189 DLNQSKFAPTVSDLLKESTQNTVSLKESSTQGVSSLFREITASSAVSILIKPQETDPC 248  
Db 181 DLNQSKFAPTVSDLLKESTQNTVSLKESSTQGVSSLFREITASSAVSILIKPQETDP- 239

Qy 249 LSCFGMSVLMFN-----AKSGRRKRXKPTXSHLMRKMSRTGLTCLKHLGQGAQR- 301  
Db 240 -----LPVSRNVGADAKCKKRRKKVTNIIISFDEDEQNSGCVFKKTPGAGESSED 294

Qy 302 -----TPTAPLS-----ISCPPLKAPSGLTPME-----SEQQLMENX 334  
Db 295 NSDRSSNVMSAFSEPGPNQSGNSKWDLSLNGBEFYQKLDVKSIDDDVDENE 354

Qy 335 FPFVFERGVVPEACEKHR-----XXRCCKXRRXRVKMLIRKEAQGPLGVARE- 382  
Db 355 DDVY-----GNSGRKRGHSESPKLENTC-----LSQWHSWAPLKVHLND 398

Qy 383 -----ATGREHLPLPAQGLSAGAAQXLRHPLPCQWRGLLOPSRCPKPPKGERDRTG 436  
Db 399 SDILFPVSGVGSFADAPLGSLNGTGPEDHVP-----DPLGRYSVEA 443

Qy 437 PRSPGSWTSVQCSQLSRPSKSSQPTVSAPVPSMTISLRQATVAMNRKDELEENR 496  
Db 444 -SSPGH-----GSPLS-----SLPSSAVPESMTISLRQATVAMNRKDELEENR 489

Qy 497 SLRNLGEMHSAAALRQEVDTLKRKVAQBERGCMKVQALASYLCYFVRRF 548  
Db 490 SLRNLGEMHSAAALRQEVDTLKRKVAQBERGCMKVQALASYLCYFVRRF 541

RESULT 3

AAE00685  
ID AAE00685 standard; protein; 784 AA.

XX AC AAE00685;  
XX DT 04-JUL-2001 (first entry)  
XX DE Human TNF receptor-associated factor binding protein, IREN-E.  
XX KW Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;  
KW TRAF2; TRAF2 binding protein; IREN; IkappaB Regulator; IREN-E;  
KW immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytostatic; tumour;  
KW AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;  
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;  
KW anorexia; anti-HIV; therapy.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Binding-site 198..388  
FT Region 326..331  
FT /note= "This motif is also present within 51 amino acid  
FT domain spanning amino acids 769-820 of NIK (Nuclear  
FT factor (NF)-kappaB Inducing Kinase) which is essential  
FT for IKK-1 (IkappaB Kinase) binding to NIK in a 2-hybrid  
FT assay and NF-kappaB activation by NIK over expression"

XX WO200116314-A1.  
XX PD 08-MAR-2001.  
XX PF 31-AUG-2000; 2000WO-IL000517.  
XX PR 02-SEP-1999; 99IL-00131719.  
XX PA (YEDA ) YEDA RES & DEV CO LTD.  
XX PI Wallach D, Malinin N, Sinha I, Leu S;  
XX WP1; 2001-281387/29.  
XX DR N-PSDB; AAD04335.  
XX PS New DNA sequence encoding Tumor Necrosis Factor receptor associated  
XX factor (TRAF) binding proteins (IREN) for treatment or prevention of  
XX pathological conditions associated with NF-kappaB induction.  
XX Claim 14; Fig 8; 118pp; English.

XX The present sequence is human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein isoform, IREN-E. A fragment of this IREN-E molecule is capable of binding to TRAF2 protein at position 225-501. The invention relates to human tumour necrosis factor (TNF) receptor-associated factor (TRAF2) binding protein designated as IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their corresponding cDNA molecules. IREN is useful for modulating/mediating the activity of transcription factor NF (Nuclear Factor)-kappaB or any other intracellular signalling activity mediated by TRAF2. IREN is useful in the prevention and treatment of a pathological condition associated with NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia, septic shock and graft-vs-host reactions. IREN also plays an important role in the control of inflammation and other non-apoptotic effects of TNF as well as in the control of apoptosis. The invention also relates to method for screening, identifying and producing a molecule capable of modulating activities mediated by IREN. IREN antibodies are useful for the purification of new proteins from different sources, including cell extracts or transformed cell lines, in addition IREN can be used in diagnostic purposes for identifying disorders related to abnormal functioning of cellular effects mediated directly by TRAF proteins

XX Sequence 784 AA;

Query Match 49.7%; Score 1537.5; DB 4; Length 784;  
Best Local Similarity 61.2%; Pred. No. 7.6e-132;  
Matches 356; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

QY 9 MSGSNXNDKROFLERLLDAVKQCIQIRFXGKKEITASDSRVTCLCAQFEAVLQHLKRS 68  
Db 1 MSGSQNDKROFLERLLDAVKQCIQIRFGGRKEITASDSRVTCLCAQFEAVLQHLKRS 60

QY 69 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVLNKHLELQRFYSLRHIASDVGRGAWLR 128  
Db 61 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVLNKHLELQRFYSLRHIASDVGRGAWLR 120

QY 129 CALNEHSLERYLHMLADRCRLSTFYEDWMSFVMDERSMPLPTWAAGLNSILFAINIDNK 188  
Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWMSFVMDERSMPLPTWAAGLNSILFAINIDNK 180

QY 189 DLNGOSKFAPTVSDLLKSTQNTVTLKKESTQVSSLFREITASSAVSILIKPEQETDPC 248  
Db 181 DLNGOSKFAPTVSDLLKSTQNTVTLKKESTQVSSLFREITASSAVSILIKPEQETDPC 239

QY 249 LSCPGMSVLMFN-----AKRSGRKRKXPTXSHLMRMKMSRTLTCLKRLHGOGRQOR- 301  
Db 240 -----LPVVSRRVNSADAKCKKXKXKVTNIIISFDDEDEQNSGDVFKTTPGAGESSED 294

QY 302 -----TTPTAPLS-----LSCPLKAPSGLTPE-----SEQOLMENX 334  
Db 295 NSDRSSVNIIMSAPFPGFNGSQSSNSWKIDSLNGEFGYQKLDVKSIDDEVDENE 354

QY 335 FPFVFERGVVPEAXCKHR-----XXRCGXKRRXRVKWLIRKEAQQPLGVARE- 382  
Db 355 DDVI-----GNSGRKRRHSHSEPKLEGNTC-----LSQMSWAPLKVLEND 398

QY 383 -----ATGREHPLPDAQGSAGAAQXLRHPLPCWQGLLQPSRCPKRGDRRTG 436  
Db 399 SDILFPVSGVGSYPADAPLGLSLENGTGPEDHVP-----DFGLRYSVEA 443

QY 437 PRSPGWSVTSVQCSQISRPKSEQPVTSASVPESMTISELQATVAMNRRKDELEENR 496  
Db 444 -SSPGH-----GSPUS-----SLIPASVPESMTISELQATVAMNRRKDELEENR 489

QY 497 SLRNLIDGEMHSAALRQEVDTLKRKVAEQEERQGMKVQALA 538  
Db 490 SLRNLIDGEMHSAALRQEVDTLKRKVAEQEERQGMKVQALA 531

RESULT 4  
AAE00684  
ID AAE00684 standard; protein; 813 AA.

XX AAE00684;  
XX  
DT 04-JUL-2001 (first entry)  
XX

Human TNF receptor-associated factor binding protein, IREN-10B.

XX Human; Tumour Necrosis Factor; TNF; TNF Receptor Associated Factor;  
KW TRAF2; TRAF2 binding protein; IkappaB Regulator; IREN-10B;  
KW immunosuppressive; nuclear factor-kappaB; NF-kappaB; cytostatic; tumour;  
KW AIDS; acquired immune deficiency syndrome; rheumatic disease; apoptosis;  
KW autoimmune disease; septic shock; graft-vs-host reaction; inflammation;  
KW anorexia; anti-HIV; therapy.

XX  
OS Homo sapiens.

XX  
FH Binding-site 198..388  
FT Location/Qualifiers  
FT /note= "TRAF2 binding region"  
FT 326..331  
FT /note= "This motif is also present within 51 amino acid  
FT domain spanning amino acids 769-820 of NIK (Nuclear  
FT factor (NF)-kappaB Inducing Kinase) which is essential  
FT for IKK-1 (IkappaB Kinase) binding to NIK in a 2-hybrid  
FT assay and NF-kappaB activation by NIK over expression"

XX WC200116314-A1.  
XX 08-MAR-2001.  
XX 31-AUG-2000; 2000WO-IL000517.  
XX 02-SEP-1999; 99IL-00131719.  
XX (YEDA ) YEDA RES & DEV CO LTD.  
XX Wallach D, Malinin N, Sinha I, Leu S;  
XX WPI; 2001-281387/29.  
XX N-PSDB; AAD04334.

New DNA sequence encoding Tumor Necrosis Factor receptor associated  
factor (TRAF) binding proteins (IREN) for treatment or prevention of  
pathological conditions associated with NF-kappaB induction.

Claim 15; Fig 7; 118pp; English.

The present sequence is human tumour necrosis factor (TNF) receptor-  
associated factor (TRAF2) binding protein isoform, IREN-10B. A fragment  
of this IREN-10B molecule is capable of binding to TRAF2 protein at  
position 225-501. The invention relates to human tumour necrosis factor  
(TNF) receptor- associated factor (TRAF2) binding protein designated as  
IREN (IkappaB Regulator), its isoforms IREN-10B, IREN-E and their  
corresponding cDNA molecules. IREN is useful for modulating/mediating the  
activity of transcription factor NF (Nuclear Factor)-kappaB or any other  
intracellular signalling activity mediated by TRAF2. IREN is useful in  
the prevention and treatment of a pathological condition associated with  
NF-kappaB induction (abnormal) e.g. AIDS (acquired immune deficiency  
syndrome), autoimmune diseases, tumours, rheumatic diseases, anorexia,  
septic shock and graft-vs-host reactions. IREN also plays an important  
role in the control of inflammation and other non-apoptotic effects of  
TNF as well as in the control of apoptosis. The invention also relates to  
method for screening, identifying and producing a molecule capable of  
modulating activities mediated by IREN. IREN antibodies are useful for  
the purification of new proteins from different sources, including cell  
extracts or transformed cell lines, in addition IREN can be used in  
diagnostic purposes for identifying disorders related to abnormal  
functioning of cellular effects mediated directly by TRAF proteins

Sequence 813 AA;

Query Match 49.7%; Score 1537.5; DB 4; Length 813;  
Best Local Similarity 61.2%; Pred. No. 8e-132;  
Matches 356; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

QY 9 MSGSNXNDKROFLERLLDAVKQCIQIRFXGKKEITASDSRVTCLCAQFEAVLQHLKRS 68  
Db 1 MSGSQNDKROFLERLLDAVKQCIQIRFGGRKEITASDSRVTCLCAQFEAVLQHLKRS 60

QY 69 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVLNKHLELQRFYSLRHIASDVGRGAWLR 128  
Db 61 RGLALTAATAAKQAAGFASKTETEPVFVYVYVKEVLNKHLELQRFYSLRHIASDVGRGAWLR 120

QY 129 CALNEHSLERYLHMLADRCRLSTFYEDWMSFVMDERSMPLPTWAAGLNSILFAINIDNK 188  
Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWMSFVMDERSMPLPTWAAGLNSILFAINIDNK 180

QY 189 DLNGOSKFAPTVSDLLKSTQNTVTLKKESTQVSSLFREITASSAVSILIKPEQETDPC 248  
Db 181 DLNGOSKFAPTVSDLLKSTQNTVTLKKESTQVSSLFREITASSAVSILIKPEQETDPC 239

QY 249 LSCPGMSVLMFN-----AKRSGRKRKXPTXSHLMRMKMSRTLTCLKRLHGOGRQOR- 301  
Db 240 -----LPVVSRRVNSADAKCKKXKXKVTNIIISFDDEDEQNSGDVFKTTPGAGESSED 294

QY 302 -----TTPTAPLS-----LSCPLKAPSGLTPE-----SEQOLMENX 334  
Db 295 NSDRSSVNIIMSAPFPGFNGSQSSNSWKIDSLNGEFGYQKLDVKSIDDEVDENE 354





PT biodiversity.  
 XX Claim 20; SEQ ID NO 58372; 103pp; English.  
 PS  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 131 AA;  
 Query Match 7.6%; Score 236; DB 4; Length 131;  
 Best Local Similarity 57.6%; Pred. No. 3.1e-13;  
 Matches 53; Conservative 5; Mismatches 16; Indels 18; Gaps 4;  
 QY 392 PDAQLGSA-----EGAAQLRHPLPCQWRGLLPSCPPKPGERDTRGPRSPGS 442  
 DB 46 PETEPASAYGPFSTETAGCAEAGNSQP-WLN-----KCPKPKGERDTRGPRSPRS 99  
 QY 443 WTSVQCSQLSRPKSSEQPVT---SASVPES 471  
 DB 100 WTSVQCSQLSRPKSSEQPVTFLCSARVHDN 131  
 RESULT 9  
 ABG62508  
 ID ABB62508 standard; protein; 720 AA.  
 AC ABB62508;  
 XX  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX Drosophila melanogaster polypeptide SEQ ID NO 14316.  
 DE  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 KW  
 XX Drosophila melanogaster.  
 OS  
 XX WO200171042-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX 23-MAR-2001; 2001WO-US009231.  
 PF  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR  
 XX 11-JUL-2000; 2000US-00614150.  
 PR  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 XX WPI; 2001-656860/75.  
 XX N-PSDB; ABL06611.  
 DR  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 XX interactions.  
 PS Disclosure; SEQ ID NO 14316; 21pp + Sequence Listing; English.  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
 CC AB872072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 720 AA;  
 Query Match 5.9%; Score 183.5; DB 4; Length 720;  
 Best Local Similarity 21.8%; Pred. No. 2.4e-07;  
 Matches 97; Conservative 61; Mismatches 162; Indels 125; Gaps 17;  
 QY 14 NXDKRQFLERLLDAVKQCOIRFXGRKEIASDSRVTCICAQFEAVLQHLKRSRGAL 73  
 DB 28 NENAREIEFE-----VRQKSLEVLGVCE-----QTSALCTTLEALFLHGLKDS----F 71  
 QY 74 TAAAIKQAGFASKTETEPVFWYVYKVLNKHQLQRFYSLRHIASDVGRGAWLRCAINE 133  
 DB 72 LSATFNVIAGDVER-RPEPSFWAPCMVFMHKQVIEQVQGLSQITSETGQCRAWVROSLNE 130  
 QY 134 HSLERYLHMLLADRCRLSTFYEDMSFVMDERSMPLTMAAGLNS-----ILFAINDNKK 189  
 DB 131 SVFSSYLNMRKNGSALSQYKRNALMRDSEG---LETAAKIMESLEAYVQFDLPVNSSL 187  
 QY 190 LNCQSKFAPTVSDI---LK-----ESTQNTVTSLLKES 218  
 DB 188 LNHWPDPYSLQLSDLWTLPALKSCPISSGVDVASSLGSDDIIAIPTPQPLQTNELFSESISNS 247  
 QY 219 TQGVSSLF--REITASSAVSILIKPEQETDPCLSCPGMSVLMPNKASGRRK-----268  
 DB 248 PFCRNENFGVPPELSQRENLDLLIQKDEMD-----VINEETGGDKAETNETAE 296  
 QY 269 RKYPTXSHLMRMKXSRTLGTCCLKRHILGGQQAQRTTTPAPLSISCPPLKAPSG-----320  
 DB 297 QEAPSGSKTRSRRKS-----NRQMDSSSFADLDLEAPSL-LPQGSNSLSNLM 342  
 QY 321 -----LTPMESEQLMENXFPVFERGVWTP-----EAXCEKHXRXR 356  
 DB 343 QTSWSGDLEATPTSPTEELTGSRF--FORSVSVSSVSLRSPTTDRCSYNALLRKHSNR 400  
 QY 357 CGKRRXRVRWKLIRKEAQGPLGVAR 381  
 DB 401 ESGAGHSEIWEKFRASASN-LNVAQ 424  
 RESULT 10  
 ABB11928  
 ID ABB11928 standard; peptide; 1019 AA.  
 XX  
 AC ABB11928;  
 XX  
 DT 11-JAN-2002 (first entry)  
 XX  
 DE Human BG33\_7 protein homologue, SEQ ID NO:2298.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibitor; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischaemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;



PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 14; SEQ ID NO 1936; 205pp; English.

XX The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 473 AA;

Query Match 5.7%; Score 175.5; DB 6; Length 473;  
 Best Local Similarity 28.5%; Pred. No. 7.2e-07;  
 Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;  
 QY 5 GXGMSGSXNDKQFLLERLLDAVKOCQIRFXGRK-----EIASDSRVTVCLCAQF 57  
 DB 14 GGGGGGKKSLSARNAVER--RNLTIVC--RFSVKTLIDRSCFTIDDSPEFNNFAIL 70  
 QY 58 EAVLQHGGLKR----SRGLALTAALK--QAAGFASKTETEPVFWYVYKVLNK---HELQR 109  
 DB 71 EQILSHRLKEISQSCRWLHLQIPLOQGVTFG--YESPRFWDYIRVACKVQNCICS 128  
 QY 110 FYSLRHIASDVGRGAWLRCAALNEHSLERYLHMLLADRCRLSTFYEDWFSVMDERSM 169  
 DB 129 IENNVSSSRKGRWIRVALMEKHLSEYISTALRDFKTRRFYEDGAILGEE--ANML 187  
 QY 170 PTMAAGLNSILFALINDKNGOSKFPATVSDLLKSTQNTVSLK--ESTQGVSSILFR 227  
 DB 188 AGMLLGLNAIDFSCFKGEGLDGS---FPAVID-----YTPYLYIOSSDSISSDEE 236  
 QY 228 EI-TASSAVSILIKPEQETDPCL 249  
 DB 237 ELRTLGSSGSESTPENVGPPFL 259

RESULT 12

ADJ70685

ID ADJ70685 standard; protein; 473 AA.

XX AC ADJ70685;

XX 06-MAY-2004 (first entry)

DE Human heat mitochondrial protein as a therapeutic target SeqID2491.

XX mitochondrial; human; screening assay; diabetes mellitus;

KW Huntington's disease; osteoarthritis;

KW Leber's hereditary optic neuropathy; LHON;

KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;

KW neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;

KW osteopathic; ophthalmological; cytostatic.

XX Homo sapiens.

XX WO2003087768-A2.

XX 23-OCT-2003.

XX 04-APR-2003; 2003WO-US010870.

XX 12-APR-2002; 2002US-0372843P.

XX 17-JUN-2002; 2002US-0389987P.

XX 20-SEP-2002; 2002US-0412418P.

XX (MITO-) MITOKOR.

XX (BUCK-) BUCK INST AGE RES.

XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;

PI Warnock DE;

XX WPI; 2003-845369/78.

XX Identifying a mitochondrial target for drug screening assays and for  
 PT treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 PT with the disease.

XX Claim 1; SEQ ID NO 2491; 180pp; English.

XX This invention relates to novel mitochondrial targets that can be used  
 CC for therapeutic intervention in treating a disease associated with  
 CC altered mitochondrial function. Specifically, it refers to a method for  
 CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins that  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, nontropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cytostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.

XX Sequence 473 AA;

Query Match 5.7%; Score 175.5; DB 7; Length 473;  
 Best Local Similarity 28.5%; Pred. No. 7.2e-07;  
 Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;  
 QY 5 GXGMSGSXNDKQFLLERLLDAVKOCQIRFXGRK-----EIASDSRVTVCLCAQF 57  
 DB 14 GGGGGGKKSLSARNAVER--RNLTIVC--RFSVKTLIDRSCFTIDDSPEFNNFAIL 70  
 QY 58 EAVLQHGGLKR----SRGLALTAALK--QAAGFASKTETEPVFWYVYKVLNK---HELQR 109  
 DB 71 EQILSHRLKEISQSCRWLHLQIPLOQGVTFG--YESPRFWDYIRVACKVQNCICS 128  
 QY 110 FYSLRHIASDVGRGAWLRCAALNEHSLERYLHMLLADRCRLSTFYEDWFSVMDERSM 169  
 DB 129 IENNVSSSRKGRWIRVALMEKHLSEYISTALRDFKTRRFYEDGAILGEE--ANML 187  
 QY 170 PTMAAGLNSILFALINDKNGOSKFPATVSDLLKSTQNTVSLK--ESTQGVSSILFR 227  
 DB 188 AGMLLGLNAIDFSCFKGEGLDGS---FPAVID-----YTPYLYIOSSDSISSDEE 236  
 QY 228 EI-TASSAVSILIKPEQETDPCL 249  
 DB 237 ELRTLGSSGSESTPENVGPPFL 259

RESULT 13

ABG21178

ID ABG21178 standard; protein; 1006 AA.

XX AC ABG21178;

XX 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #21169.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.





Qy 80 QAAGFASKT---ETEPVFWYVVKVNLKHELORFYSLRHIAADVGRGAWLRCAALNEHSL 136  
Db 2002 QASGYAPKSFHVEDTPVCF-----SRNSSLSLSID-----SEDDLL 2038  
Qy 137 ERYLHMLLADRLSTTFYEDWSFVMDERSMLPTMAAGLSILFAINIDNKDL---NGQ 193  
Db 2039 QECISSAMPKKRPSR-----LKSESEKQSRKVGGI--LAEDITLIDKDLQRPDSE 2088  
Qy 194 SKFAPTVDL---LKSTQNVTSLLKSTQGVSSLFREITASSAVSILK-----240  
Db 2089 HAFSPGSENFOWKAIQEGANSIVSLHQAAAAACLSQASSSDSLSLSKSGISLSPF 2148  
Qy 241 ---PEQETDCLSCPGKSVLMP-----NAX 262  
Db 2149 HLTPDQBEKPTSNKGRILPKGBKSTLEAKKIESENKGIKGGKVKYKSLITGKIRNSE 2208  
Qy 263 RSGRRKRXKPTXSHLMWRKMSRTLTCLKRLHGGRAQRTTPTAPLSICPPLKAPSGLT 322  
Db 2209 ISSQWKQPLPTNMFISR--GRTM-----IHI-FGLRNSSSSTSPVKKGPPLKTPASKS 2260  
Qy 323 PMESEQQLMENXFFVFERGVWVPEACEKHXKXRCGXKRRXRVWVKLIRKEAQGPGVARE 382  
Db 2261 PSEG-----PGATTSPRCTKPAKSELSPIRQTSQISGSKNGSSRS 2302  
Qy 383 ATGREHLPPLPAQLGSAEGAAOXLRHLPQOWRGLLOPSR---CPPRKPGERDRTRGP-- 437  
Db 2303 GS-RDSTPSRPTQ-----QPLSRPMQSPGRNISPGRNGISFPNKLSQLPRTSPST 2353  
Qy 438 ---RSPG---SMTSVQSGSLSPR--KSSEQPVTASVPESMTISE-LRQATVAM-MN 486  
Db 2354 ASTKSSSGKWSYISP--GROLSQNLTKQASLSKNASSIPRSSASKGLNQMSGNGSN 2411  
Qy 487 RKDELEENRSLRNLIDGE-MEHSAAALRQEV-----DTLKRKXVABQEE-----528  
Db 2412 KKVLSRMSSTKSSGSDSSEPALVRQSTFIKEAPSTLRRKLEESASFESLSPSSRP 2471  
Qy 529 ---RQMKVQALASYLCYFVRFXPHVTRTWWR-----NGGR-EKSNSSXS 571  
Db 2472 DSPTRSOAQTPLVSPSLPMSLSTHPSVQAGGWRKLPNLSPTTEYDNGRPTKRHDARS 2531  
Qy 572 HLSS 575  
Db 2532 HSES 2535  
RESULT 15  
ID ABG21174 standard; protein; 388 AA.  
XX AC ABG21174;  
XX DT 18-FEB-2002 (first entry)  
XX DE Novel human diagnostic protein #21165.  
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX OS Homo sapiens.  
XX FN WO200175067-A2.  
XX PD 11-OCT-2001.  
XX PF 30-MAR-2001; 2001WO-US008631.  
XX PR 31-MAR-2000; 2000US-00540217.  
XX PR 23-AUG-2000; 2000US-00649167.  
XX PA (HYSE-) HYSEQ INC.  
XX FI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.  
XX N-PSDB; AAS85361.  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX Claim 20; SEQ ID NO 51533; 103pp; English.  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 388 AA;  
Query Match 5.0%; Score 154; DB 4; Length 388;  
Best Local Similarity 25.7%; Pred. No. 5.1e-05;  
Matches 53; Conservative 29; Mismatches 76; Indels 48; Gaps 6;  
Qy 95 WY---YVKEVNLKHELORFYSLRHIAADVGRGAWLRCAALNEHSLRYLHMLLADRLCRLS 151  
Db 53 WWRMDWTPRLPSRHHIIELEHLTFVNDVGRGAWLRCAALNEHSLRYLHMLLADRLCRLS 112  
Qy 152 TFYEDWSFVMDERSMLPTMAAGLSILFAINIDNKDLN-----GQSKFAPTVDLL 204  
Db 113 EYQPTALLRDABEGEPFLSFLQGLMSLSFELSYKSAIINLNEWTLTPLALSGLCP-LSELD 171  
Qy 205 KESTQNVTSLLKSTQGV-----SSLFREITASSAVSIL 238  
Db 172 PLSTSGAELQKESLDSHSGSEDIKVHSHGHKTRRNOKPTASSLSLDTASSSQLSCS 231  
Qy 239 IKPEQETDPCLSCPGMSVLMFNAKRS 264  
Db 232 L-----NSDCL-----LQENGSKS 246  
Search completed: June 7, 2005, 11:57:19  
Job time : 97.0732 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:43:10 ; Search time 20.8699 Seconds  
(without alignments)  
2784.626 Million cell updates/sec

Title: US-09-155-676B-2  
Perfect score: 3093  
Sequence: 1 XTGPXGXMGSGXNKKRQF.....FLMTFHVCEPINCFLSLKKK 604

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score.	Query Match	Length	DB ID	Description
1	178.5	5.8	624	2 T27096	hypothetical prote
2	160	5.2	2845	2 I49505	adenomatous polypo
3	136.5	4.4	2843	1 RBHUP	adenomatous polypo
4	130.5	4.2	364	2 T43467	hypothetical prote
5	123	4.0	2272	2 C54972	voltage-dependent
6	122	3.9	2222	2 A37490	probable VPI/AB13
7	120.5	3.9	780	2 G84708	actin-interacting
8	118.5	3.8	1033	2 T37715	protein H3923.1
9	118	3.8	1245	2 G89287	probable serine/th
10	113.5	3.7	1192	2 T18611	centromere protein
11	113	3.7	905	2 A54654	microtubule-associ
12	112.5	3.6	5327	2 T13564	hypothetical prote
13	111.5	3.6	792	2 T29187	hypothetical prote
14	111	3.6	553	2 T45872	hypothetical prote
15	111	3.6	1291	2 T17242	beta spectrin, bet
16	111	3.6	2388	2 J60271	smooth muscle myos
17	110.5	3.6	1938	2 JC5421	smooth muscle myos
18	110.5	3.6	1972	2 JC5420	protein M01A8.2
19	109.5	3.5	893	2 G88551	nuclear mitotic ap
20	109.5	3.5	2101	2 A42184	hypothetical prote
21	109	3.5	992	2 T46337	kinesin motor prot
22	109	3.5	1459	2 T30196	giantin - human
23	108.5	3.5	3225	2 I52300	Bassoon protein -
24	107.5	3.5	3942	2 T42730	hypothetical prote
25	106	3.4	623	2 T16167	homeotic protein C
26	106	3.4	975	2 S33121	basal transcriptio
27	106	3.4	1469	2 T09219	calcium channel pr
28	106	3.4	2223	2 A47447	microtubule-associ
29	106	3.4	2774	2 A43359	

30	105.5	3.4	699	2 T01029	hypothetical prote
31	105.5	3.4	840	1 S69204	pheromone response
32	105	3.4	1323	2 T00037	hypothetical prote
33	104.5	3.4	651	2 T15624	hypothetical prote
34	104.5	3.4	858	2 S15762	neurofilament trip
35	104	3.4	1008	2 S15762	hypothetical prote
36	104	3.4	1200	2 AE2304	neurofilament prot
37	103	3.3	1171	2 T00380	KIAA0637 protein -
38	103	3.3	3924	2 S37431	ankyrin 2, neuro
39	102.5	3.3	878	2 T17245	hypothetical prote
40	102	3.3	2441	2 S19161	CREB-binding prote
41	101.5	3.3	1332	1 I48314	homeotic protein C
42	101.5	3.3	3259	1 A56539	giantin - human
43	101	3.3	237	1 S47630	translation elonga
44	101	3.3	1616	2 I37183	gene APXL protein
45	101	3.3	2007	1 B43402	myosin heavy chain

ALIGNMENTS

RESULT 1

T27096  
hypothetical protein Y51H1A.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T27096  
R;Smye, R.  
submitted to the EMBL Data Library, October 1998  
A;Reference number: Z20309  
A;Accession: T27096  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-624 <WIL>  
A;Cross-references: UNIPROT:Q9XWJ7; EMBL:AL032644; PIDN:CAA21666.1; GSPDB:GN00020; CESP:  
C;Experimental source: clone Y51H1A  
C;Genetics:  
A;Gene: CESP:Y51H1A.2  
A;Map position: 2  
A;Introns: 6/3; 57/2; 92/3; 151/2; 268/1; 414/3; 535/1  
C;Superfamily: Caenorhabditis elegans hypothetical protein Y51H1A.2

Query Match	5.8%	Score 178.5;	DB 2;	Length 624;
Best Local Similarity	26.6%;	Pred. No. 2.1e-05;		
Matches	71;	Conservative 44;	Mismatches 117;	Indels 35; Gaps 9;
Qy	29	VKQCQIRFXGRKE-----IASDSRSVTC-----	LCAQPEAVLQGLKESRGLAL	73
Db	24	LKQEKIRADLQKELDNVVKAAIASPHSKRENVPSEITQNLCSNTEAIFHGLRDPFFLKG	83	
Qy	74	TAAAIKQAGFASKTETETEPVFWYVYKVNKHELRQFYSLRHIASDVGRGRAMLRCALNE	133	
Db	84	TRYA-----KYPEPNFPVSKFSRHSITSQIACLGQIRSEIGKSPAWIRIVLNE	133	
Qy	134	HSLEYVHMLADRCRLSTPYEDWSP---VMDERSMPLPTMAAGLSILFPAINDKDL	190	
Db	134	NALQGYLDLLAAEATAIQFYSDDAFLRLSDGQSERIRGLLKLPLSLSPISATNSFL	193	
Qy	191	NGQSKAPTVDLLKSTQNTVSLKSTQVSLFTEITASSAVSVILIKPEQETDPCLS	250	
Db	194	NTWTPPLIILAGLMDGQPLKVGTLKARPKNPKAHLTEEI-AIPDALV-PEEDHD--IG	249	
Qy	251	CPGMSVLMPNAKRS-GRKKKXPTXSH	276	
Db	250	SP--SYLEKRRRLRSRPIKSENDDH	274	

RESULT 2

I49505  
adenomatous polyposis coli protein - mouse  
N;Alternate names: APC  
C;Species: Mus musculus (house mouse)  
C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004

C;Accession: I49505  
R;Su, L.: 668-670, 1992  
A;Title: Multiple intestinal neoplasia caused by a mutation in the murine homolog of the A;Reference number: I49505; MUID:92263101; PMID:1350108  
A;Accession: I49505  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-2845 <RES>  
A;Cross-references: UNIPROT:O61315; GB:M88127; NID:g191991; PIDN:AAB59632.1; PID:g191992  
C;Superfamily: adenomatous polyposis coli protein  
Query Match 5.2%; Score 160; DB 2; Length 2845;  
Best Local Similarity 20.2%; Pred. No. 0.0037;  
Matches 122; Conservative 80; Mismatches 224; Indels 178; Gaps 27;  
QY 80 QAAGPASKT---ETEPVFWYVKEVNLKHELQRFVSLRHASDVGRGAWLRCAINHESL 136  
DB 2002 QASGYAPKSFHVEDTPVCF-----SRNSLSLSID-----SEDDL 2038  
QY 137 ERYLHMLLADRCRLSTFYEDWDFVMDERSMLPTMAAGLNSILFAINIDNKDL---NGQ 193  
DB 2039 QECISSAMPKKRPSPR-----LKSESEKQSPKVGGI--LAEDLTLDLKLQRPDSE 2088  
QY 194 SKFAPTVSDL---LKESTQNTVSLKESTQGVSSLPREITASSAVSLIK----- 240  
DB 2089 HAFSPGSENFMDKAIQEGANSIVSSLHQAAAAACLSRQSSDSLSLSKSGISLSP 2148  
QY 241 ----PEQETDPLSCPGMSVMP-----NAX 262  
DB 2149 HLTPDQEEKFTNKGPRILKPKGEKSTLEAKIESENKGIKGGKVKYKSLITGKIRNS 2208  
QY 263 RSGRRKXKPTXSHLMWKMRTLTGTLCKRHLCQGRQRTTTPALISCPPLKAPSGLT 322  
DB 2209 ISSQMKQPLTNMPSISR--GRTM-----IHI-PGLRNSSTSPVSKGPKPLTPASKS 2260  
QY 323 PMESEQOLMENXFPVFERGVWVPEACEKHXRCXKGRXRVWKLIRKEAQGLGVARE 382  
DB 2261 PSEG-----PGATTSRGTGPKAGKSLSPITQTSIGSNKGSRS 2302  
QY 383 ATGREHLPLDQAQGAAGAAQXLRHPLPCQWRGLLQPSR---CPPRKPGERDRTRGP-- 437  
DB 2303 GS-RDSTPSRPTQ-----QPLSRPMQSPGRNSISPCGRNGISPPNKLSQLPRTSSPT 2353  
QY 438 ---RSPG-----SWTSVOCGQLSRPR--KSSEQPVTSASVPESMTISE-LRQATVAM-WN 486  
DB 2354 ASTKSSQSGKMSYTSR--GRQLSQNLTKQASLSKNASSIPRSESASKGLNONGNNGSN 2411  
QY 487 RKQELEENSLRLNDGE-MEHSAAALRQEV-----DTLKKKVAEQEE----- 528  
DB 2412 KKVLSRMSSTKSGSSESSESRPALVRQSTFIKEAPSLRLKLEESASFESLSPSSRP 2471  
QY 529 ----RQGMKVQALASYLCYFVRFPXPHVRTWWR-----NGGR-EKSNSSXXS 571  
DB 2472 DSPTRSOAQTPVLSPLPDMSLSTHSPVQAGWRKLPNLSPTIEYNDGRPTKRHDIA 2531  
QY 572 HLSS 575  
DB 2532 HSES 2535  
RESULT 3  
RBUHAP  
adenomatous polyposis coli protein - human  
N;Alternate names: polyposis coli locus protein DP2.5; tumor suppressor APC  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text\_change 09-Jul-2004  
C;Accession: A37261; B39658; A44928; A49319; I54271  
R;Kinzler, K.W.; Nilbert, M.C.; Su, L.K.; Vogelstein, B.; Bryan, T.M.; Levy, D.B.; Smith  
Chui, S.F.; Horii, A.; Ando, H.; Miyoshi, Y.; Miki, Y.; Nishisho, I.; Nakamura, Y.  
Science 253, 661-665, 1991  
A;Title: Identification of FAP locus genes from chromosome 5q21.  
A;Reference number: A37261; MUID:91335210; PMID:1651562

A;Accession: A37261  
A;Molecule type: mRNA  
A;Residues: 1-2843 <KIN>  
A;Cross-references: UNIPROT:P25054; GB:M74088; NID:g182396; PIDN:AA03586.1; PID:g182397  
R;Joslyn, G.; Carlson, M.; Thilveris, A.; Albertsen, H.; Gelbert, L.; Samowitz, W.; Grod  
arlington, J.; McPherson, J.; Wasmuth, J.; Le Paslier, D.; Abderrahim, H.; Cohen, D.; Le  
Cell 66, 601-613, 1991  
A;Title: Identification of deletion mutations and three new genes at the familial polypos  
A;Reference number: A39658; MUID:91330307; PMID:1678319  
A;Accession: B39658  
A;Molecule type: DNA  
A;Residues: 1-183, 'L', '185-969, 'N', '971-1308, 'G', '1310-1324, 'SS', '1326, 'HSTLE', '1332-1354, 'P',  
A;Cross-references: GB:M73548; NID:g190163; PIDN:AA60354.1; PID:g190164  
R;Miki, Y.; Nishisho, I.; Horii, A.; Miyoshi, Y.; Utsunomiya, J.; Kinzler, K.W.; Vogelst  
Cancer Res. 52, 643-645, 1992  
A;Title: Disruption of the APC gene by a retrotransposal insertion of L1 sequence in a c  
A;Reference number: A44928; MUID:92119623; PMID:1310068  
A;Accession: A44928  
A;Molecule type: DNA  
A;Residues: 1506-1525 <MIK>  
A;Cross-references: GB:S78214; NID:g243541; PIDN:AAB21145.1; PID:g243542  
A;Note: Sequence extracted from NCBI backbone (NCBIN:78214, NCBIP:78218)  
R;Spurio, L.; Olschwang, S.; Groden, J.; Robertson, M.; Samowitz, W.; Joslyn, G.; Gelber  
Cell 75, 951-957, 1993  
A;Title: Alleles of the APC gene: an attenuated form of familial polyposis.  
A;Reference number: A49319; MUID:94073973; PMID:8252630  
A;Accession: A49319  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 'G', '143-171, 'P', '173-179 <SP1>  
A;Cross-references: GB:S67787; NID:g461061; PIDN:AAD13997.1; PID:g4261697  
R;Lambertz, S.; Ballhausen, W.G.  
Hum. Genet. 90, 650-652, 1993  
A;Title: Identification of an alternative 5' untranslated region of the adenomatous poly  
A;Reference number: 154271; MUID:93186137; PMID:8383094  
A;Accession: 154271  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-4 <LAM>  
A;Cross-references: GB:S56365; NID:g266243; PIDN:AAD14918.1; PID:g4262770  
C;GeneticS:  
A;Gene: GDB:APC  
A;Cross-references: GDB:119682; OMIM:175100  
A;Map position: 5q21.5q22  
A;Note: mutations of this gene can result in familial adenomatous polyposis or sporadic  
C;Superfamily: adenomatous polyposis coli protein  
C;Keywords: cancer; familial adenomatous polyposis; tumor suppressor  
F;1-730/Domain: leucine-rich <NTD>  
F;7-72/Region: coli #status predicted  
F;185-227/Region: coli #status predicted  
F;731-2832/Domain: serine-rich <CTD>  
F;1131-1156/Region: acidic  
F;1558-1577/Region: acidic  
F;1866-1893/Region: highly charged  
Query Match 4.4%; Score 136.5; DB 1; Length 2843;  
Best Local Similarity 21.9%; Pred. No. 0.21;  
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;  
QY 156 DWSFVMDERSMLPTMAAGLNSILFAINIDNKDLNGQSKFAPTVSDLLKESTQNTVSL 215  
DB 2100 DMKAQEE-----GANSIVSSLSL-----QAAAAACLRSQSSDSLSL- 2138  
QY 216 KESTQGVSSLPREITASSAVSLIKPEQETDPLSCPGMSVLMNAK----- 262  
DB 2139 -KSGISLSPFH-----LTPDQEEKFTNKGPRILKPKGEKSTLETKKIESBSK 2186  
QY 263 ---RSGR-----KRKXPTXSHLMWKMRTLTGTLCKRHLCQGRQRTT 302  
DB 2187 GKGGKKVYKSLITGKVRNSISQMKQPLQANMPSISRGRTM-----IHI-PGVRNS 2240  
QY 303 TPTAPLISCPPLKAPSGLTPMESQQLMENXFPVFERGVWVPEACEKHXRCXKGR 362

Db 2241 SSTSPVSKGKPLKTPASKSPSEGTATTS-----PRGA-KPSVKSE----- 2281

Qy 363 XRVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAEG-----AAQXLRHPLPCQWRGL 417

Db 2282 --LSPVARQTSQ--IGSSKAPSR-----GSRDSTPSRPAQQLSRPIQSPGRNS 2328

Qy 418 LQPSR---CPRKRGDRTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464

Db 2329 ISPRNGISPPNKLSQLPRTSSP--STASTKSSGGRKMSYTPGQMSQQNLTKQTGLSKN 2387

Qy 465 SASYPESWTISE-LQATVAM-MNRKDEL-----EENRSLRNLIDGEMEHSAAAL 512

Db 2388 ASSIPRESASAKGLNQNNGCANKVELSRMSSTKSSGSESDRSRPVL---VRQSTFI 2444

Qy 513 ROEYD-TLKRKVAE 525

Db 2445 KEAPSPTLRKLEE 2458

RESULT 4

T43467

hypothetical protein DKFZp434A1727.1 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004

C:Accession: T43467

R:Bioecker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: 22513

A:Accession: T43467

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-364 <AAA>

A:Cross-references: UNIPROT:Q9UF50; EMBL:AL133657

A:Experimental source: adult testis; clone DKFZp434A1727

C:Genetics:

A:Note: DKFZp434A1727.1

C:Superfamily: human hypothetical protein DKFZp434A1727.1

Query Match 4.2%; Score 130.5; DB 2; Length 364;

Best Local Similarity 25.0%; Pred. No. 0.04;

Matches 53; Conservative 36; Mismatches 98; Indels 25; Gaps 7;

Qy 41 EIADSDSRVTCLOCAFEAVLQHLKRSRGLALTAALKQAAGPASKTETEPVWYVKE 100

Db 49 EPIDDSSEEFVNFVAAILEQLSHRFK-----ACAPAGPVSWFSS--DGQRGFWDYIRL 99

Qy 101 VLNK---HELQRFYSLRHIAISDVGRGAWLRCALENEHSLRYLHMLADRCLSTFFYEDW 157

Db 100 ACSKVPNNCVSSIIENMENISTARAKGAWIRKVALMEKRMSEYITLARDTRTTRRFYDSG 159

Qy 158 SFVNDSESRSMPLTMAAGLSILPAINDKDLNGQSKFAPTVSDLLKESTQNVTSLLKE 217

Db 160 AIMLRDE-ATILTQMLGLSADISFCLKGEVLGDKT---PVVID-----YTPYLK- 206

Qy 218 STQGVSSLFREITASSAVSILIKPEQETDPCLPCLSCPGMSVLMPNKRSRRK--- 270

Db 207 FTQSYDYLTDEERHSAESSTSDNSPEHPVL 238

RESULT 5

C54972

voltage-dependent calcium channel alpha 1E - mouse

C:Species: Mus musculus (house mouse)

C>Date: 12-Apr-1995 #sequence\_revision 12-Apr-1995 #text\_change 09-Jul-2004

C:Accession: C54972

R:Williams, M.E.; Marubio, L.M.; Deal, C.R.; Hans, M.; Brust, P.F.; Philipson, L.H.; Mili J. Biol. Chem. 269, 22347-22357, 1994

A:Title: Structure and functional characterization of neuronal alpha-1E calcium channel

A:Reference number: A54972; MUID:94350992; PMID:8071363

A:Accession: C54972

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: mRNA

A:Residues: 1-2272 <WIL>

A:Cross-references: UNIPROT:Q61290; GB:L29346; NID:9522330; PIDN:AAA59206.1; PID:9522331

A:Note: authors translated the codon AGG for residue 788 as Lys, and CCT for residue 886

C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 4.0%; Score 123; DB 2; Length 2272;

Best Local Similarity 20.8%; Pred. No. 1.6;

Matches 90; Conservative 58; Mismatches 181; Indels 104; Gaps 19;

Qy 155 EDWSFWDESRSMPLTMAAGLSILPAINDKDLNGQSKFAPTVSDLLKESTQNVTSL 214

Db 658 EDWNEVMYNGIRSQ--GGVSSGMSAIYFIVL---TLFGNTYLLNVFLAIVDNLANAQEL 713

Qy 215 LKSTQGVSSLFREITASSAVSILIKPEQETDPCLPCLSCPGMSVLMPNKRSRRK--- 270

Db 714 TKDE-QEIEEAFNQKHA-----LQKAKVSP-MSAPN-----MPSIERDRRRHHMSMW 760

Qy 271 XPTXSHLMKMKSRRTLG-----TCLKRHLGQG-----RAQRTTPTAPLSICPPLKAPS 319

Db 761 EPRSSHLRERRRRHHMSVWEQRTSQLRRHMQMSQEAALNKEEAPPMPLN----PLNPLS 816

Qy 320 GLTPMSESEQQLMENXFPVFERGVVWPEAKCEKHXRCGXKRRXRVWKLIRKEAQGLGV 379

Db 817 PLNPLNAPHSLYRPRPIEGIALGLGLEKEEERISRG-----SLKGDIGG 863

Qy 380 ARSATGREHLPLPDAQLGSAEGAAQXLRHP--LPCQWRGLLP-----SR 422

Db 864 LTSALDNQRSPL---SLGK-----REPPWLPKRSCHGNCDPQOBAGGETVVTFRDR 912

Qy 423 CRRKRGDRTRGPRSPGWSVQCGSQLSRPKSSEQPVTSASVPESMTISBLRQATV 482

Db 913 ARHRQSQRRRHRVRTEG-----KOSASASRSASQSE-----RSLDEGVSVGEKEHEP 963

Qy 483 AMWNRKDE--LEENR-----LRNLIDGEMEHSAAALRQGVDTLKRKVA 524

Db 964 HSHRSRKEPTIHEERTQDRLRTNSLMVPRGSLVGLADENETPLVQPELEVGKDAAL 1023

Qy 525 EQEERQGMKVOAL 537

Db 1024 TEQEAEGSSQAL 1036

RESULT 6

A37490

voltage-dependent calcium channel alpha 1E - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A37490

R:Soong, T.W.; Stea, A.; Hodson, C.D.; Dubel, S.J.; Vincent, S.R.; Snutch, T.P. Science 260, 1133-1136, 1993

A:Title: Structure and functional expression of a member of the low voltage-activated ca

A:Reference number: A37490; MUID:93262464; PMID:8388125

A:Accession: A37490

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-2222 <SOO>

A:Cross-references: UNIPROT:Q07652; GB:L15453; NID:g310082; PIDN:AAA40855.1; PID:g310083

A:Experimental source: brain

A:Note: sequence extracted from NCBI backbone (NCBI:P:132101)

C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 3.9%; Score 122; DB 2; Length 2222;

Best Local Similarity 20.8%; Pred. No. 1.8;

Matches 90; Conservative 58; Mismatches 181; Indels 104; Gaps 19;

Qy 155 EDWSFWDESRSMPLTMAAGLSILPAINDKDLNGQSKFAPTVSDLLKESTQNVTSL 214

Db 608 EDWNEVMYNGIRSQ--GGVSSGMSAIYFIVL---TLFGNTYLLNVFLAIVDNLANAQEL 663

Qy 215 LKSTQGVSSLFREITASSAVSILIKPEQETDPCLPCLSCPGMSVLMPNKRSRRK--- 270

Db 664 TKDE-QEIEEAFNQKHA-----LQKAKVSP-MSAPN-----MPSIERDRRRHHMSMW 710

Qy 271 XPTXSHLMKMKSRRTLG-----TCLKRHLGQG-----RAQRTTPTAPLSICPPLKAPS 319

Db 711 EPRSHLRERRHHMSVWEQRTSQLRRHMQMSQEAALKEEAPPMPLN-----PLNPLS 766

Qy 320 GLTPMESEQOQLENKFPVFERGVVPEAXCEKHXRRXRCGXRRXRVWKLIRKEAQGLGV 379

Db 767 PLNPLNAHPSLYRRPPIEGLAGLGLKECEERISRG-----SLKGDIGG 813

Qy 380 AREATGREHLPLPDAQLGSAEAAQXLRHP--LPCWRGLLOPS-----R 422

Db 814 LTVSLDNGKRSPL---SLGK-----REPPWLPKRSCHGNCDPQOETGGGETVVTPEDR 862

Qy 423 CPKPRGGERDRGRGSPGSMVTSVQCGSLRPRKSPGSPVTSASVPESMTISLRQATV 482

Db 863 ARHROSQRSHRRVRTEKES-----ASASRSASQSE-----RSLDEGVSDKEKEHEP 913

Qy 483 AMNRKDE--LEENRS-----LRNLDGMEHSAALRQEVDTLKRKA 524

Db 914 QSHRSKEPTIHEERTQDLRTNSLMVPRGSLVGLDGAETPLVQPPLEVGKDAAL 973

Qy 525 EOEERQGMKVQAL 537

Db 974 TEQEAEGSGSEQAL 986

RESULT 7

G84708

probable VP1/AB13 family regulatory protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: G84708

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: G84708

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-780 <STO>

A;Cross-references: UNIPROT:O04346; GB:AE002093; NID:g1946371; PIDN:AAB63089.1; GSPDB:GN

C;Genetics:

A;Gene: At2g30470

A;Map position: 2

Query Match 3.9%; Score 120.5; DB 2; Length 780;

Best Local Similarity 21.8%; Pred. No. 0.61;

Matches 122; Conservative 75; Mismatches 243; Indels 119; Gaps 30;

Qy 38 GRKEIASDSRVYTCQAQFEAVLQGLKRSR-GLALTAALAKQAAGFASKTETE----- 91

Db 202 GMKHISPSQSNVHCASN---ILQ---KPSRPAISTPPVASKSAQARIGRPVVEGRG 255

Qy 92 ---PVFW---YVKEV-----LNKHELQRPYSLRHIASVGR-GRAWLRCALNEH----- 134

Db 256 HLLPRYPKPYTDEVQVQISGNLNLVPLPEKTLASDAGRIGRLVLPKACAEAYFPPI 314

Qy 135 SLERYLHMLAD-RCRLSTYEDWSFVM-----DEERSMLPTMAAGLSILF-----A 184

Db 315 SQSEGIPLKIQDVRGR-----EWTQFRYPWNNSRMVLEGVTPCISMMQLQADVT 368

Qy 185 IDNKDLNGQ---SKFAPTVSDLLKSTQNVTSLLKESTQGVSSLPFREITASSAVSILIK 240

Db 369 FSRVDPGKILMSGRKKAAGDQGCGLTNGTSTEDTSSGVTENPPSSGSCISLIPK 428

Qy 241 -----PEQETDPCLSCPGMSVLMPNAKSGRRKRRKXKPTXSHLMWRKMSRGLTCLXR--- 292

Db 429 ELNGMPEN-----LNSETNGRIGDDPT--RVKCKRTRTIGAKNRLLL 471

Qy 293 HLGQGRQARTPTAPLISCPPLKAPSGLTPMSEQOLME-NXFPVVF-ERGVVWVPEAXCE 350

Db 472 HSESMELRUTWEAAQDLRPSVKPTIVVIE-EQEIEYDEBPVFGKTIYTKPSGE 530

Qy 351 KHXRCXGKRRXR-----VWKLIRKEAQGLGVARE-----ATGREH 388

Db 531 QERWATCDDCSKWRRLPVDALLSFKWTCIDNVWDVSRSCSPAESLKELENVLKVGREH 590

Qy 389 LPLPDAQLGSAEAAQXLRHPPLPCWRGLLOPSRCPPR--KPGERDRTGRSPSGSWTSV 446

Db 591 ---KERTGESAQAKSQE---PCGLDALASAAVLGDTTIGEPEVATTTTTPRHRACGSCI 644

Qy 447 QCGSQLSRPR-KSSEQPVTSASVPESMTISLRQATVAMNRRKDELEENRSRLNLDGE 505

Db 645 VC---IQPPSGKGRHKPTCGCTVCSTV---KRFKTLMMRRKKQLERDVTAAEDKKKD 698

Qy 506 ME-----HSAALRQEVDTLK 520

Db 699 MELAESDKSKEKEVNTAR 717

RESULT 8

T37715

actin-interacting homolog SPAC15A10.16 [imported] - fission yeast (Schizosaccharomyces P

C:Species: Schizosaccharomyces pombe

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 15-Sep-2000

C:Accession: T37715

R;Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, June 1997

A;Reference number: Z21738

A;Accession: T37715

A;Status: translated from GB/EMBL/DBD

A;Molecule type: DNA

A;Residues: 1-1033 <MUR>

A;Cross-references: EMBL:Z97208; PIDN:CA10112.1; GSPDB:GN00066; SPDB:SPAC15A10.16

A;Experimental source: strain 972h; cosmid c15A10

C;Genetics:

A;Gene: SPAC15E1.01; SPDB:SPAC15A10.16

A;Map position: 1

Query Match 3.8%; Score 118.5; DB 2; Length 1033;

Best Local Similarity 21.7%; Pred. No. 1.2;

Matches 95; Conservative 67; Mismatches 152; Indels 123; Gaps 23;

Qy 161 MDEERSMLPTMAAGLN-----SILFAINIDNK-----DLNGQSKFAPTVS 201

Db 177 LNQSRN-LPONTREANDNVNRNRSRGSSGSSISFRNPDSRKQYEITTSASLSAFGAF 235

Qy 202 DLLKE-----STQNVTSLLK--ESTQGVSSLPFREITASSAVSILIKPEQETDPCLS 252

Db 236 DALQTNALERRASKRLSHVMVKDQNNEGSQNLNITVES-VRGFLRDSKPDNIMDSP 294

Qy 253 G-----MSVLMPNAKSGRRKXKPTXSHLMWRKMS--RTLGTCLKHLQGRQARTTP 304

Db 295 SPKVSKRPSIVRODSHDSNKLRR-PTINTSFDKRFSPKLTTRTSLTKSLDPG-----TP 348

Qy 305 TAPLSICPPL-KAPSGLTPMSEQOLMENFFPVERGVVWVPEAXCKHXXXRCXKRX 363

Db 349 T---SLKSPSLKSPSSFVQKD-----VYSRS-----NSLRISOANRS 383

Qy 364 RVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAEAAQXLRHPPLPCWRGLLOPSRC 423

Db 384 NVF-----PGATDVTRSVS--DHRILSSSTINDGEVAP-----PLP----- 417

Qy 424 PRKGERDRT-RGPRSPGSMVTSVQCGSQLSRPR-KSSEQPV-----TSASVPES 471

Db 418 -----QSRSTISSPNSLATVLPSSPTPILLPGRSSTLSVNNKQFNADGGSTLNSPS 471

Qy 472 MTISLRQATVAMNRRKDELEENRSRLNLDGEME--HSAALRQEVDTLKRKA----- 524

Db 472 IRETEYAAAPKLEIADEVETDATSQRELLERQIKAESSEDTSEISLQGLSLPQVS 531

Qy 525 ---EOEERQGMKVQALAS 539

Db 532 STQOEIQPSSSVPEAAS 548



Query Match	3.6%	Score 112.5	DB 2	Length 5327
Best Local Similarity	18.7%	Pred. No. 29		
Matches	80	Conservative 64	Mismatches 176	Indels 107
Gaps	17			

  

QY	188	KDLNGQKFPAPTVDLLKESQTQNTSL-----LKESQTGVSSLPREITAS	232
Db	2936	KDEAKESRRSEVAESVKPSSDKDATSAPPSKEHSRPSVGLSKD--EGDKTTSRRVSVA	2993
QY	233	SAV-----SILL-----KPEQETDPC--LSCPQMSVLMEAKSRGRKRKXPTXSHLMWRK	281
Db	2994	DSIKDEKSLIVSOAGRPSEASESLKDAAPQGETSRPESVTESVKQCKSPVASKASRP	3053
QY	282	MSRTLIGTCLKHHLGQGAQRTTPTAFLSTSCPLKAPS---GLTPM-----ESE	327

3054	AS	--VAGNAKASDAESKEQKFE	---	380
328	QY	QOOLMENXFPFVERGVWVPEAXCEKHRRXCGKRRXRVRWKLIRKEAQOPLGVA	---	380
3105	Db	RESVAEOFFPLVSKVEVSPASVAESVKDEAEKKEBS	---PLMSKEASRSPASVAGSVKDEA	3161
381	QY	---REATGREHL	---PLPDAQLSGAEQAQXLRHPLPCQWRGLLOPSCRCPKPERDR	433
3162	Db	EKSKEESRRESVAEKSPLPSKEASRPASVAESV	-----KDEADKSK	32020
434	QY	TEGPSPSGSWTSVQCGSQLSRP	-----RKSSQPVTSASVPES-MTISELRQAT	481
3203	Db	ESRRESGAEKSPLASKEASRPASVAESIKDEAEKSKEESRRESVAEKSPLPSKEASRPT	3262	
482	QY	VAMNRKDELEE-ENRSLRN	-----LLDGEHMSAALROEVDTLKRKVAEQEEORQMCK	533
3263	Db	SVAKSVKDEAEKSKEESRSDRVAEKSPLASKEASRPASVAESVQDEAEKSKEESRESVA	3322	
534	QY	VOALASY	540	
3323	Db	EKSPLAY	3329	

RESULT 13  
T29187  
hypothetical protein C5C3.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29187  
R:Woessne, J.; Stellyes, L.  
submitted to the EMBL Data Library, April 1996  
A:Description: The sequence of C. elegans cosmid C5C3.  
A:Reference number: Z20585  
A:Accession: T29187  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Molecule type: DNA  
A:Residues: 1-792 <WOE>  
A:Cross-references: EMBL:U53335; PIDN:AAA96170.1; GSPDB:GN00022; CESP:C5C3.3  
A:Original source: strain Bristol N2; clone C5C3

[illegible]



Db 264 WYDFLKLIVRYHCHGNRPLLKIALTAVFGISIIIVAYRLIVKTRRRGGNIKKKPAK 323  
Qy 240 -----KPEQETDPCLSCP---GMSVLPNNAKSGRRKRXPTX 274  
Db 324 PLKRITPNNSDEKLKMKSPPKPQLOATSDILQPLPSNNAMTVPESKESANR---SPQS 380  
Qy 275 SHL-MWRKMSSTLTCLKRLHGGRA-----QRTTPTAPLSISCPLKAPSLTP 323  
Db 381 SKINLVNRGGRTVNTDLCDCSEARAIPKCDKLTDFQSTNSEQSPADAIASAVAPTCLTG 440  
Qy 324 MESSQQLME-----NXPPVPERGVMP-EAXCEKHXKXRCGXKRRXRV--WKLI----- 369  
Db 441 SOSSEKQIEIKMPSNRFELQDSFKFAEADIPPEERKSGTEQSTTEPTKLAGSQSPS 500  
Qy 370 -RKEAQGLGVAREATGREHLPLD--AQLGSABGAQXLRHPLPCQWRGLLOPSRCPPR 426  
Db 501 ERQEIQQMPSNRFELQDSFKFAEADIPPEERKSGTEQSTTEPTKLAGSQSPS 543  
Qy 427 KPGDRDRTRGPRSPGWSVVOGSQLSRPKSKSQPVTSASVPMSITSELURQATVAMN 486  
Db 544 -----PCN-----NRRNSETSQVDTISPVPKLVGTQ-----SPS 573  
Qy 487 RKDELEENRSLNLLDGEHSAALRQEVDTLKRKVAEQERQGMKVQALA 538  
Db 574 ERQEIQQMPSNRFELQDSFKFAEADIPPEERKSGTEQSTTEPTKLAGSQSPS 625

RESULT 14  
T45872  
hypothetical protein F4P12.20 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 09-Jul-2004  
C;Accession: T45872  
R;Blocher, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat, M.  
submitted to the Protein Sequence Database, January 2000  
A;Reference number: 223016  
A;Accession: T45872  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-553 <BLO>  
A;Cross-references: UNIPROT:Q9LFI5; EMBL:AL132966  
A;Experimental source: cultivar Columbia; BAC clone F4P12  
C;Genetics:  
A;Map position: 3  
A;Introns: 35/1; 108/1; 197/1; 209/3; 231/3; 281/1; 491/3  
A;Note: F4P12.20

Query Match 3.6%; Score 111; DB 2; Length 553;  
Best Local Similarity 20.1%; Pred. No. 2;  
Matches 93; Conservative 66; Mismatches 177; Indels 126; Gaps 22;  
Qy 153 FYEDMSFVMDERSMLPTMAAGLNSILFAINIDKOLNGSKFAPTVSDLLKESTQNTV 212  
Db 103 FTSAGVLEPELSSM-----SSNHKSGK-KALPTILEDINKSTESIS 145  
Qy 213 SLKKESTQGVSS---LFREITA---SSAVSILIKPEQE-----TDPCLSCPGMSVLMPN 260  
Db 146 TFQSDCTVENSQEFVLDFEDVPASIQRSKTSVATPKSNVLRATDVAIS-PTSSITVDVT 204  
Qy 261 AKRGRRKRXKPTXSHLMRMKMSLTGLCKRHLHGGRAQRTTPTAPLSISCPPLKAPSG 320  
Db 205 ATQGTKTSKGGP-----RNPSRV-----QPGKATK-QPVATRGLSTISKPPNG 248  
Qy 321 LTPMESQQLMENFPFVPERGVWVPEAXCEKHXKXRCGXKRRXRVWKLIRKEAQP--LGV 379  
Db 249 LSKVRPLSTSTN-----RSLDISKTQEQKNKSLPAG-----KEPLGPRISM 291  
Qy 380 AREATGREHLPLPDAQLGSABGAQXLRHPLPCQWRGL---LQPSRCPPRKP----- 428  
Db 292 SRR--KPVLPKPGVPFKSSRSRSDASKNEMTSSCSLSCASASSASHKPSIDSIKKK 349  
Qy 429 -----GERDRTRG-----PRSPGWSVVOGSQLSRPKSKSQPVTSASVPESM 472

Db 350 NDSSRSLUSSQPLANRSTSRGIMQPRIPPOQT-----NKTSPKPLSSSVP-TAGSISDYS 403  
Qy 473 TISELROATVAMNRRKDELEENRSLNLLDGEHSAALRQEVDTLKRKVAEQERQGM 532  
Db 404 SSSRASSETSKMANGKQTVSRKVPAN--DNTVQTVKPLKNSKDT---SVQADAKEGT 458  
Qy 533 K-VOALASYLCYFVRREXPVHRTWWRNGGREKSNSSXXSHL 573  
Db 459 KRVSAT-----NGGLVPSASAKPSGL 479

RESULT 15  
T17242  
hypothetical protein DKFZp586B1417.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T17242  
R;Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A;Reference number: Z18722  
A;Accession: T17242  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1291 <KOE>  
A;Cross-references: UNIPROT:Q9UFV0; EMBL:AL117448  
A;Experimental source: adult uterus; clone DKFZp586B1417  
C;Genetics:  
A;Note: DKFZp586B1417.1

Query Match 3.6%; Score 111; DB 2; Length 1291;  
Best Local Similarity 20.4%; Pred. No. 6;  
Matches 39; Conservative 41; Mismatches 71; Indels 40; Gaps 7;  
Qy 7 GXMGSXNXXDKROFLLERLLDAVK--QCQIRFXKGRKEIASDSRVVTCICAQFEAVLQHG 64  
Db 757 GLLEKCRNKTMRMLVEKMGREAVELGHGVNITGVBE-----NTLIASLCOLLERISHG 811  
Qy 65 LKRSRGLA-----LTAATAIKQAAGFASKTETE-----PVFWYVVK 99  
Db 812 LQVKGKSAWLSHLLHYQDNQRKLTSGSL-STSGILLDSERRKSDASSLMPPL----- 864  
Qy 100 EVLNKHELQRFYSLRHASDVGRGAWLRCAALNEHSLERYLHMLADRCLSTFYEDWSF 159  
Db 865 RISLIQDMRHQIONIGEIKTDVGKARAWRLSMKLLSRHLKQLLSDHELTKKLYKRYAF 924  
Qy 160 V-MDEERSM 169  
Db 925 LRCDDKEQFL 935

Search completed: June 7, 2005, 12:03:26  
Job time : 24.2033 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:40 ; Search time 88.3902 Seconds  
(without alignments)  
3499.208 Million cell updates/sec

Title: US-09-155-676B-2  
Perfect score: 3093  
Sequence: 1 XTPGKGXMSGXNXDKRQF.....FLWTFHVCBPINCQFSLKXK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1196	38.7	375	2 Q9HA26	Q9ha26 homo sapien
2	400.5	12.9	249	2 Q8IU14	Q8iu14 homo sapien
3	358.5	11.6	373	2 Q8TEQ0	Q8teq0 homo sapien
4	329	10.6	476	2 Q9D3S3	Q9d3s3 mus musculus
5	308.5	10.0	249	2 Q8K050	Q8k050 mus musculus
6	302	9.8	520	2 Q9VKJ1	Q9vkj1 drosophila
7	257	8.3	1056	2 Q6P2R5	Q6p2r5 homo sapien
8	257	8.3	1058	2 Q9Y4G2	Q9y4g2 homo sapien
9	251	8.1	1074	2 Q7TS11	Q7ts11 mus musculus
10	190.5	6.2	1087	2 Q80TQ5	Q80tq5 mus musculus
11	189	6.1	1095	2 Q8IWE5	Q8iwe5 homo sapien
12	183.5	5.9	720	2 Q9W293	Q9w293 drosophila
13	183	5.9	1020	2 Q94Z98	Q94z98 homo sapien
14	178.5	5.8	604	2 Q9XWJ7	Q9xwj7 caenorhabdi
15	175.5	5.7	473	2 Q96NLO	Q96nlo homo sapien
16	172	5.6	408	2 Q6PDC0	Q6pdc0 mus musculus
17	171	5.5	387	2 Q8TBG7	Q8tbg7 homo sapien
18	167	5.4	496	2 Q7QES7	Q7qes7 anopheles g
19	166	5.4	708	2 Q7I593	Q7i593 homo sapien
20	160	5.2	310	2 Q96P31	Q96p31 xenopus tro
21	160	5.2	712	2 Q8BIJ7	Q8bij7 mus musculus
22	160	5.2	2845	1 APC MOUSE	Q61315 mus musculus
23	159.5	5.2	1010	1 Y226 HUMAN	Q92622 homo sapien
24	157.5	5.1	961	2 Q6PD22	Q6pd22 mus musculus
25	155	5.0	329	2 Q8IWM5	Q8iwm5 homo sapien
26	151	4.9	600	1 RUF1 HUMAN	Q96t51 homo sapien
27	149	4.8	590	2 Q96P51	Q96p51 homo sapien
28	149	4.8	606	2 Q8WXA3	Q8wx3 homo sapien
29	149	4.8	619	2 Q9P1Z1	Q9p1z1 homo sapien
30	148	4.8	600	2 Q9EPM6	Q9epm6 mus musculus
31	147.5	4.8	941	2 Q8BMP7	Q8bmp7 mus musculus

#### RESULT 1

ID	Q9HA26	PRELIMINARY;	PRT;	375 AA.
AC	Q9HA26;			
DT	01-MAR-2001 (TrEMBLrel. 16, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Hypothetical protein FLJ12363.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Mammary gland;			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,			
RA	Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,			
RA	Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,			
RA	Abe K., Kamiyama K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,			
RA	Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,			
RA	Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y., Ishida S.,			
RA	Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,			
RA	Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,			
RA	Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,			
RA	Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,			
RA	Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takenoto M., Kawakami B.,			
RA	Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,			
RA	Fujinori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,			
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,			
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,			
RA	Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,			
RA	Mizushima-Shigano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,			
RA	Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,			
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isoqai T., Sugano S.,			
RT	"Complete sequencing and characterization of 21,243 full-length human			
RT	cDNAs."			
RL	Nat. Genet. 36:40-45(2004).			
RL	EMBL; AK022425; BAB14033.1; --			
DR	GO; GO:0003700; F:transcription factor activity; IEA.			
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.			
DR	InterPro; IPR002197; HTH_Fis.			
DR	InterPro; IPR004012; Run.			
DR	Pfam; PF02759; RUN; 1.			
DR	PRINTS; PR01590; HTHFIS.			

Q8by22 mus musculus  
Q8r4c2 mus musculus  
Q69zh1 mus musculus  
Q9cav9 mus musculus  
Q6nww8 mus musculus  
Q8bl21 mus musculus  
Q88fk7 xenopus lae  
Q8vdl0 mus musculus  
Q8bkq4 mus musculus  
Q80y95 mus musculus  
Q08576 mus musculus  
Q7zt23 brachydanio  
Q8n4u6 homo sapien  
Q9d394 mus musculus

#### ALIGNMENTS

```
DR SMART; SMO0593; RUN: 1.
DR PROSITE; PS50826; RUN: 1.
SQ SEQUENCE 375 AA; 41861 MW; F7DFBCBB4297CD95 CRC64;

Query Match 38.7%; Score 1196; DB 2; Length 375;
Best Local Similarity 82.2%; Pred. No. 1e-73;
Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;

QY 9 MSGSNXDKRQFLRLRLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFEAVLOHGLKRS 68
Db 1 MSGSQNDKROFLRLRLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFEAVLOHGLKRS 60

QY 69 RGLALTAARAIKQAAQAFASKTETEPVFVYVYVKEVLNKHLELQRFYSLRHIASDVGRGAWLR 128
Db 61 RGLALTAARAIKQAAQAFASKTETEPVFVYVYVKEVLNKHLELQRFYSLRHIASDVGRGAWLR 120

QY 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSMPLPTWAAGLNSILFAINIDNK 188
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSMPLPTWAAGLNSILFAINIDNK 180

QY 189 DLNGOSKFAPTVSDLLKKESTQNTVTSLLKKESTQGVSSLPREITASSAVSILIKPEQETDPC 248
Db 181 DLNGOSKFAPTVSDLLKKESTQNTVTSLLKKESTQGVSSLPREITASSAVSILIKPEQETDPC 239

QY 249 LSCPGMSVLMFN-----AKRSGRKRXKPTXSHLMKMKSRITLGTCLKRHLGQORA 299
Db 240 -----LPVVSNRVNSADAKCKKRRKKKKVNTNIISFDDEDEQNSGDVFKKTPGAGES 291

RESULT 2
Q8IU14 PRELIMINARY; PRT; 249 AA.
AC Q8IU14;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE FLJ12363 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Primary B-Cells;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Primary B-Cells;
RX Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041583; AAH41583.1; -.
SQ SEQUENCE 249 AA; 27224 MW; 5F665B3F6BCE7FEB CRC64;

Query Match 12.9%; Score 400.5; DB 2; Length 249;
Best Local Similarity 64.8%; Pred. No. 1.7e-19;
Matches 94; Conservative 11; Mismatches 27; Indels 13; Gaps 4;

QY 161 MDERSMPLPTWAAGLNSILFAINIDNKDINGOSKFAPTVSDLLKKESTQNTVTSLLKKESTQ 220
Db 1 MDERSMPLPTWAAGLNSILFAINIDNKDINGOSKFAPTVSDLLKKESTQNTVTSLLKKESTQ 59

QY 221 GVSLSLPREITASSAVSILIKPEQETDPCISCPGMSVLMFN-----AKRSGRKRXKPTXSH 276
Db 60 GVSLSLPREITASSAVSILIKPEQETDPCISCPGMSVLMFN-----LPVVSNRVNSADAKCKKRRKKKQVTNI 113

QY 277 LMM--RMSSTLTGCLKRHLGQORA 299
Db 114 ISFDDEDEQNSGDVFKKTPGAGES 138

RESULT 3
Q8TEQ0 PRELIMINARY; PRT; 373 AA.
AC Q8TEQ0;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE FLJ00143 protein (Fragment).
GN Name=FLJ00143;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074072; BAB84898.1; -.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR Pfam; PF00787; PX; 1.
DR SMART; SM00312; PX; 1.
DR PROSITE; PS50195; PX; 1.
FT NON TER 1
SQ SEQUENCE 373 AA; 42541 MW; 2C34370BCCFABF7B CRC64;

Query Match 11.6%; Score 358.5; DB 2; Length 373;
Best Local Similarity 81.0%; Pred. No. 2.2e-16;
Matches 81; Conservative 1; Mismatches 5; Indels 13; Gaps 2;

QY 439 SPGSWTSVQCGSQLSRPRKSKSEQPVTSAPESMTISELROATVAMNKRKDELEENRSL 498
Db 5 SPGH-----GSLPS-----SLLPSASVPESMTISELROATVAMNKRKDELEENRSL 51

QY 499 RNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKVOALA 538
Db 52 RNLLDGEHSAALRQEVDTLKRKVAEQEERQGMKVOALA 91

RESULT 4
Q9D3S3 PRELIMINARY; PRT; 476 AA.
AC Q9D3S3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4933437K13 product:hypothetical PX (Bem1/NCF1/P13K)
DE domain containing protein, full insert sequence (SNX29).
DE Name=4933437K13Rik; Synonyms=Snx29;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```



Mon Jun 13 13:27:11 2005

```

499 RNLLDGEHSAALRQEVDTLTKRKVAEQEERQGMKVQALA 538
|||||
50 RNLLDGEHSAALRQEVDALERRKVTQEQRHATKVQALA 89
|||||

RESULT 6
Q9VK31 PRELIMINARY; PRT; 520 AA.
AC Q9VK31.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CG5439-PA (LD35592p).
GN ORFNames=CG5439;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari P., Brottier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter A., Chandra I.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fogle C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiland T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RA "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
SEQUENCE FROM N.A.
MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomic perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
SEQUENCE FROM N.A.
MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
SEQUENCE FROM N.A.
FLYBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A.
FLYBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
[7]
SEQUENCE FROM N.A.
STRAIN-Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[8]
EMBL; AE003639; AAF53251.1; -.
DR EMBL; AY095047; AAM11375.1; -.
DR FLYBase; FBgn0032476; CG5439.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001683; PX.
DR InterPro; IPR004012; Run.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00195; PX; 1.
DR PROSITE; PS50826; RUN; 1.
DR PROSITE; PS00678; WD_REPEATS; 1; UNKNOWN 1.
SQ SEQUENCE 520 AA; 59256 MW; 2BA4F6EB39AC605 CRC64;
Query Match 9.8%; Score 302; DB 2; Length 520;
Best Local Similarity 29.4%; Pred. No. 2.4e-12;
Matches 75; Conservative 59; Mismatches 99; Indels 22; Gaps 7;
QY 9 MSGSXXNKKRQFLLELLDAVKNQCOIRFXGRKEIASDSRVTCICAQFEAVLQHLKRS 68
DB 22 LAGSTFGQRREDIFRRLOQSAHISQKXFSG-KELATERDESVOELCESLEELMSGLKRS 80
QY 69 RGLALTA---IKQAAGFASKTETEPVFWYVVKVINKHELFQRYSLRHASD 119
DB 81 ACTSFSASFTQNNQWVSGNAGGS--NNDATWFCQTHLTHERQRYMDLQIWTN 139
QY 120 VGRGRAWLRCALNEHSLERYLHMLADRCRLSTFTYEDWSFVMDERSMPLTMAAGLNSI 179
DB 140 VGRGRAFRATLNEKRLHSHVLTWLSDEQLHRFTWPSLLNDAAKKLPFVDSLSDV 199
QY 180 LFAINDNKLNGQSKFAPTVSDDLKE-----STQNTVSLKESQTGVSSLFREITASSA 234
```

```

Db 200 L FALNVDTELNAPRRSTPSVA--VKGEPIFTTSPVPVVGQRKRPGL-AVERDIECVSS 256
Qy 235 VSIL---IKPQETD 246
Db 257 TEDLGLGALKPIESVE 271

RESULT 7
Q6P2R5 Q6P2R5 PRELIMINARY; PRT; 1056 AA.
AC Q6P2R5;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Pleckstrin homology domain containing, family M (With RUN domain)
member 1.
GN Name=PLEKHM1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J.J., Helton E., Kettman M., Madan A.C., Rodriques S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strauberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC064361; AAH64361.1;
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR004012; Run.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS50081; DAG PE BIND_DOM_2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50826; RUN; 1.
SQ SEQUENCE 1056 AA; 117442 MW; 24004093CFA89079 CRC64;

Query Match 8.38; Score 257; DB 2; Length 1056;
Best Local Similarity 23.88; Pred. No. 7.3e-09;
Matches 122; Conservative 68; Mismatches 183; Indels 140; Gaps 21;

```

```

Qy 23 ERLIDAVKQCIQIRPKRKETASDSRVTCLCAQFEAVLOHGLKRSGLTAATAAIKQAA 82
Db 20 KLVGVSKALQKQVSLDTVTYVSEDGDANTMCSALEAVFIHGLHAKHRAEAGGKRKSA 79
Qy 83 GFASKTETEPFVYVVKVKNKHELQRFYSIRHTASDVGRGRAWLRCAALNEHSLEYLHM 142
Db 80 --HOKPLPQVPFWPLLKAVTHKHIISELEHLTFVNTDVGRCAWLRLALNDGLMECYLKL 137
Qy 143 LLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLNSILFAINIDNKDLN-----GQSK 195
Db 138 LLQEQARLHEVYQPTALLRDAERGEPLSLFQGLTSLSFELSYKSAILNEWTLTPLALSG 197
Qy 196 PAPTVDLLKESTQNTVSLKKESTQGVSSLF-----REITASS----- 233
Db 198 LCP-LSLDELSTSGAELQKESLDSTSHSGSEDIIEVHSHGKHRRNQKLTASSLSLDT 256
Qy 234 AVSILIKPEQETDPL-----SCPGMSVLMNPKNRGRKRXKPTXSHL 277
Db 257 ASSQSLSCSLNSDCLQENGSKSPDHCEPMSGDS-DLGTANAEDSDRSLOE----- 308
Qy 278 MMRKMSRTLTCLKRLHGLGQRAQT---TPTAPLS-----ISCPPLKAPSGLTPME 325
Db 309 VLLEFSKAQVNSVPTN---GLSQETETPTQASISLHGLNTSTYLHC---EAPAEPLPAQ 362
Qy 326 SEQOLMENKFPVFERGVWPEAXCEKHXRCGKRRXRVWKLIRKEAQGLGVAR--EA 383
Db 363 AASGTQD-----GVHVQEP-----RPQAPSPDLQQPVES 392
Qy 384 TGREHLPLDPA-----QLGSAEAGRAQLRHLPLPCOMRGLQLQSPRCPKRGDRTRG 436
Db 393 TSGQQ---PSTVSETAREVGQNGQLQKAQHD-----GAGLKLVVSS 432
Qy 437 PRSP--GSWTSVQCGSQLSRKSPKSSQEPVTSAS 467
Db 433 PTFPKNKSWSIS---EDDFYRP--SREQPLESAS 460

RESULT 8
Q9Y4G2 Q9Y4G2 PRELIMINARY; PRT; 1058 AA.
AC Q9Y4G2;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE KIAA0356 protein (Fragment).
GN Name=KIAA0356;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Yamakawa H., Ohara O.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AB002354; BAA20813.2;
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR004012; Run.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF02759; RUN; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00593; RUN; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
DR PROSITE; PS50081; DAG PE BIND_DOM_2; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR PROSITE; PS50826; RUN; 1.
FT NON TER 1
SQ SEQUENCE 1058 AA; 117670 MW; CD1AD1869C8C9D1 CRC64;

```

RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E., Schnerich A., Schein J.E., Jones S.J., Marra M.A.
RA	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL	[2]
RL	SEQUENCE FROM N.A.
RP	STRAIN=C57BL/6; TISSUE=Brain;
RC	Strausberg R.;
RA	Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL	-1- SIMILARITY: Contains 2 PH domains.
CC	EMBL: BC053079; AAH53079.1; -.
DR	GO: GO:0007242; P:intracellular signaling cascade; IEA.
DR	InterPro: IPR000345; CytC_heme_BS.
DR	InterPro: IPR02219; DAG_PE-Bind.
DR	InterPro: IPR001849; PH.
DR	InterPro: IPR004012; Run.
DR	Pfam: PF00169; PH; 1.
DR	Pfam: PF02759; RUN; 1.
DR	SMART: SM00109; C1; 1.
DR	SMART: SM00233; PH; 2.
DR	SMART: SM00593; RUN; 1.
DR	PROSITE: PS00190; CYTOCHROME C; UNKNOWN 2.
DR	PROSITE: PS50081; DAG_PE_BIND_DOM_2; 1.
DR	PROSITE: PS50003; PH DOMAIN; 2.
DR	PROSITE: PS50826; RUN; 1.
DR	SEQUENCE 1074 AA; 118533 MW; 56EDA05514E4B9F9 CRC64;
SQ	
Query Match	8.1%; Score 251; DB 2; Length 1074;
Best Local Similarity	24.5%; Pred.No.1.9e-08;
Matches 132; Conservative	58; Mismatches 209; Indels 140; Gaps 21;
Qy	19 QFLERLLDAVKCOIRFXKKEITASDSKRVTCCLCAQFEAVLQHGKRSRGLATAAAI 78
Db	15 QVKKKLVGSVKALQKHVSIDTIVTSEDGDANTMCSALEAVFIHGLH-----AKHI 66
Qy	79 KOAAGFASKTET-----EPVFWYVYKEVLNKHQLQRFYSLRHTASDVGRCAWLRCALN 132
Db	67 RAEAGGKRKKTHQKALPQPVFWPLLKAITHRHITVSDLEHLVFTINTDVGRCRAWLRALN 126
Qy	133 EHSLELYHMLLADRCILSTFYEDWSFWMDEERSMPLTMAAGLNSILFAINIDNKOLN- 191
Db	127 DGLMECYLKLILQEPARKCEYYQPTALLRDAEAEFLLSFLOGLTSLSFELSYKAILNE 186
Qy	192 -----GOSKFAPTVSDLLKESTQNVTSLKKESTQGVSSLF-----RE 228
Db	187 WTLTPLSLSLGLCP-LSELDPLTTSGAEIQKESLDSISHSSEDEIEVQSHGKHIRNRK 245
Qy	229 ITASS-----AVSILIKPEQETDCLSCPGNSVLMPNAKRSGRKRXKPTXSHLMRWKWS 283
Db	246 LTASSLSLDTASSQLSCSLNSDSL-----LQENGPKS-----PDHSEEPMSYDS 291
Qy	284 RTLGTCRLRHLGQGRQAQTTTPTAPL-SISCPPLKAPSGLTPTWESQQLMENKFPVFERGV 342
Db	292 -----DLGMANTDDPDRLSQEVLSEFBSKAQVNSAPSSGPNQ--EPDTPMFQTP 338
Qy	343 WYPEAXCKEHRXXRCGXKRRXRVWKLIRKEAQGLGVAREAT-----GREHLPLP----- 392
Db	339 SLHSLATSTH-----LHFEQSEELFPAKSSGTSGGGHKQLLLQETPD 382
Qy	393 DAQLGSAG-GAAQXLRHLPLCQ-----WRGLLQPSRCP-----PRKP----- 428
Db	393 EKQLGTAGPAQSTSDQOPSSPVGGAAGCGSPWKALEYGRVGPKLVSPTSPPKGSW 442
Qy	429 -GERDTRGRPSR-----GSWTSVQCCSOLSR-----PKS-SEQPVTASVP 469
Db	443 ISEDDFCRPPQFALKSAAGLCTSPVDPTPESRAALHGFSGQPRKSCSLGALDKACVP 501
RESULT 10	
Q80TQ5	

RESULT 10  
Q80TQ5

```

Query Match      8.3%; Score 257; DB 2; Length 1058;
Best Local Similarity 23.8%; Pred. No. 7.3e-09;
Matches 122; Conservative 68; Mismatches 183; Indels 140; Gaps 21;

QY 23 ERLDAVQCOQIRFXGKEIASDSRVTCLCAQFAVLQHGKSRGALTAQAIAKQA 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 22 KKLGVSKALQKQVSLDTVTVTSEDGDANTMCSALEAVFIHGHAHAKHRAEAGGRKKESA 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 83 GFASKTETEPFVFWYVKEVLNKHQLORFYSLRHIAASDVGRGAWLRCAINHSRLRYLHM 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 --HOKPLPQPVFWPLAKAVTHKIIISELEHLTFWTDVGRCAWLALNDGLMECYLKL 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 143 LLADRCRLSTFYEDSWFMDDEERSMLPTWAAGLSILFAINDKDLN-----GOSK 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 LLQEQARLHEYQPTALLRDAEAGEFLSLPQGLTSLSPELSYKSAILNEMWTLTPLALSG 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 FAPVSDLLKESTQNTVSLKESTQGVSSLF-----REITASS----- 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 LCP-LSELDPLSTGAELQRKESLDSISHSGSIEDIEVHSHGKIRRNQKLTAASLSLDT 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 234 AVSILIKPEQETDPCLE-----SCPGMSVILMPNAKSGRRKRKXTXSHL 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 ASSSQLSCLNSDSCLLQENGSKSPDHCEBPMSCDS-DLGTANAEDSDRSLOE----- 310
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 278 MMRKMSRVTGLTKRHLGQGRAQRT---TTTAPLS-----ISCPPLKAPSGLTWME 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 311 VLLEFSKAQVNSVPTN---GLSQBTEIPTQASLSLHGLMTSTYLHC---EAPAPPLPAQ 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 326 SFQQLMENKFFVFERGVWVPEACXKHKRXXRCGKRRXRVWKLIRKEAQGPLGVAR--EA 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 AASGTQD-----GVHVQEP-----RQAPSLDLQOPVES 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 TGREHLPPLPDA-----OLGASGAAQXLRHLPQCMWRGLLOPSRCPPKPCRDRTRG 436
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 395 TSGQQ---PSSVTSETAREVGQNGLOKAAQHD-----GAGLKLVVSS 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 437 PRSP--GSWTSVQCSQLSRPRKSEQPVTSAS 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 PTSPKNKSWIS---EDDFYRP--SREQPLESAS 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
Q77SI1
ID Q77SI1 PRELIMINARY; PRT; 1074 AA.
AC Q77SI1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE pleckstrin homology domain containing, family M (With RUN domain)
DE member 1.
GN Name=Plekhl1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP STRAIN=C57BL/6 N.A.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038959;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA whiting M., Madan A., Young A.C., Shveychenko Y., Bouffard G.G.,

```





Db 66 CGGGGCGGSAPEGEVDRILENISUSVKLQSYFAACEDEIPAINRHHKVLORLCEHL 125  
 Qy 58 EAVLQHLKRGRLGLALTAAAIKQAAFGASKTETBPFWYVYVYKVLNKHQLRFSYLRHIA 117  
 Db 126 DHALLYGLQ-----DLSSGYVWLTVHFTREAIKQIVLQHVA 163  
 Qy 118 SDVGRGAWLRCALNEHSLERYLHMLLADRCRLSTFYEDWFSVWDESSMLPTMAAGLN 177  
 Db 164 TNLGRSAWLVLALNENSLESYLRFOENLGLLHKYVKNALVCSHDHLLTLFLFLVSGLE 223  
 Qy 178 SILFPAINTDKNLQCSKFAPTVSDLLK-ESTQNTVSLKSTQGVSSL-----FREITAS 232  
 Db 224 FIRFELDLAPYLD----LAPMPDYKPYKPVLLDFEDRLPSVSGSLSLNSFNSTVST 279  
 Qy 233 SAV--SILIKPEQET----DPCLSGPGMSVLPNNA-----KRGRRKRXKPTXSHL 277  
 Db 280 NLEWDSALAPSSDYDFGDFVPAVPS----VPSTDWEDGDLTDTVSGPRS----TASDL 331  
 Qy 278 MWRKMSRTLTGTLKRLHGGRAORTTPTAPISISCPPLK--APSGLTWSEQQLMENXF 335  
 Db 332 TSSKASTRSPT-----QRQNP-----FNEEPAETVSSDTPVTHTTSQEKE--- 372  
 Qy 336 PVFERGVVWPEAXCEKHXRXCGKXRXRWKLIKREAQGLGVAREATGREHLPPLDAQ 395  
 Db 373 ---EQAQDPPDPACTELEVIRVTKK-----KIGKK-----KSRSDSEASPL---- 412  
 Qy 396 LGSABGAQXLRHPLFCWGRGLQSPRCPKPRGERDRTRGPRSGWTSVQCGSQLSRP 455  
 Db 413 -----HPACSQ-----KKC--AKQGDGSRNGSPSLGRDSP----DTMLASP 448  
 Qy 456 RKSSQEPVTSASVPSSMTISELRQATVAMNRKD-ELSEENRSLNLLD---GEMEHS 509  
 Db 449 QBEQEGP--SSTTES--SERSEPLIPMKDTSMERLQGLPSKVIDQLNGQLDQDPS 500

RESULT 12  
 Q9W293 PRELIMINARY; PRT; 720 AA.  
 AC Q9W293;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE CG6613-PA.  
 GN ORFNAMES=CG6613;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.B.D.,  
 RA Brandon R.C., Rogers Y.H., Blazie R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abriel J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burkis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fofler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celnik S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celnik S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 a genome perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bertencourt B.R., Celnik S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003456; AAF46799.1; --  
 DR FlyBase; FBgn0034694; CG6613.  
 DR GO; GO:000151; C:ubiquitin ligase complex; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0007243; P:intracellular signaling cascade; IEA.  
 DR GO; GO:0016567; P:protein ubiquitination; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR002219; DAG\_FE-bind.  
 DR InterPro; IPR004012; Run.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF02759; RUN; 1.  
 DR SMART; SM00184; RING; 1.

[illegible]

```

[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RA Smye R.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL032644; CAA21666.2; -.
DR PIR: T27096; T27096.
DR WormBase: WBGene00013093; Y51H1A.2.
DR WormPep: Y51H1A.2; CB32758.
DR InterPro: IPR004012; Run.
DR Pfam: PF02759; RUN; 1.
DR SMART: SM00593; RUN; 1.
DR PROSITE: PS50826; RUN; 1.
DR Hypothetical protein.
KW SEQUENCE 604 AA; 66549 MW; AB317F7F2B63201 CRC64;

Query Match      5.8%; Score 178.5; DB 2; Length 604;
Best Local Similarity 26.6%; Pred. No. 0.00085;
Matches 71; Conservative 44; Mismatches 117; Indels 35; Gaps 9;

QY 29 VKQCQIRFXGRK-----IASDSDSRVTC-----LCAQFEAVLQHLKRSRGLAL 73
DB 24 LKQEKIRADLQKELDNVVAATASHSKKENVPSBITQNLNSIEAIFHGLRDPFPLKG 83
QY 74 TRAAIKQAAGFASKTETEFVFWYVYKVELNKHLEQRFYSLRHIAISDVGRGRAWLRCALNE 133
DB 84 TRYA-----KYPEPFWFVSKFSHRGITSQIACLGQIRSEIGKSRAMIRIVLNE 133
QY 134 HSLERVLHMLLADRCRLSLTFYEDWSP---VMBEERSMLPTMAAGLNSILFAININDKDL 190
DB 134 NALGOYLDLAAEATAIQOYSDAFRLRLSDGQSERIRGLKPLSLPISAATNSFL 193
QY 191 NGOSKEAPTVDLLKESTQNTVTSLLKESTQGVSSLFREITASAVSILIKPQETDPCLS 250
DB 194 NTWPTPLIAGLMDGQPLKVGTLKARNPKPAHLTEEI-AIPADALV-PEEDHD--IG 249
QY 251 CPGNSVLMVNAKRS-GRKRRKXTXSH 276
DB 250 SP--SYLEKRRRRALSRPIKSENDH 274

RESULT 15
Q96NLO PRELIMINARY; PRT; 473 AA.
AC Q96NLO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ30671.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX Pubmed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita K., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba T., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Tanase T.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togliya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshihara Y., Matsumura H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,

```

```

RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yanashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
DR EMBL: AK055233; BAB70882.1; -.
DR InterPro: IPR004012; Run.
DR Pfam: PF02759; RUN; 1.
DR SMART: SM00593; RUN; 1.
DR PROSITE: PS50826; RUN; 1.
SQ SEQUENCE 473 AA; 52813 MW; 1AC96C90379EEFB8 CRC64;

Query Match      5.7%; Score 175.5; DB 2; Length 473;
Best Local Similarity 28.5%; Pred. No. 0.00099;
Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;

QY 5 GXGMSGSXNKKRQFLLERLLDAVQCQIRFXGRK-----EIASDSDSRVTCCLCAQF 57
DB 14 GGGGGGKKLSARNAVER-RNLIITVC--RFSVKTLIDRSCPETIDDSPEFNFAAIL 70
QY 58 EAVLQHLGKLR-----SRGLATAAAIK-QAAGFASKTETEFVFWYVYKVELNK---HELQR 109
DB 71 EQILSHRLKEISOSCRWLAHLQIPLOGQVTFG--YESPRSFWDYIRVACRKVSQCICS 128
QY 110 FYSLEHTASDVGRGAWLRCAALNEHSLERYLHMLLADRCRLSTFYEDWSPWMDDEERSMML 169
DB 129 IENNENVSRRAKRAWIRVALMEKHLSEYISTALRDFKTRFRFYEDGAIVLGE-ANML 187
QY 170 PTMAAGLNSILFAINIDNKDNGOSKFAPTVSDLLKESTQNTVTSLLK--ESTQGVSLFR 227
DB 188 AGMLLGLNAIDFSFCLKGEGLDGS---FPAVID-----YTPYLKYOSSDSISDDEE 236
QY 228 EI-TASSAVSILIKPEQETDPCL 249
DB 237 ELRTLGSGSGSESTPENVGPPFL 259

Search completed: June 7, 2005, 12:02:13
Job time : 92.3902 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:51:20 ; Search time 23.6321 Seconds  
(without alignments)  
1907.914 Million cell updates/sec

Title: US-09-155-676B-2  
Perfect score: 3093  
Sequence: 1 XTGPXKXMGSGXNXXKQRF.....FLWTFHVCEPINCFSHLKXK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	295.5	9.6	270	4	US-09-270-767-43983
2	151	4.9	704	4	US-09-370-838-191
3	151	4.9	704	4	US-09-854-133-191
4	136.5	4.4	2842	1	US-07-741-940-7
5	136.5	4.4	2842	1	US-08-289-548A-7
6	136.5	4.4	2842	1	US-08-452-654-7
7	136.5	4.4	2842	1	US-08-449-731-7
8	136.5	4.4	2843	1	US-07-741-940-2
9	136.5	4.4	2843	1	US-08-289-548A-2
10	136.5	4.4	2843	1	US-08-452-654-2
11	136.5	4.4	2843	1	US-08-452-655B-7
12	136.5	4.4	2843	1	US-08-452-655B-7
13	136.5	4.4	2843	2	US-08-370-235A-2
14	136.5	4.4	2843	3	US-08-450-582-2
15	136.5	4.4	2843	3	US-08-450-582-7
16	136.5	4.4	2843	4	US-08-449-731-2
17	136.5	4.4	2843	4	US-10-092-138A-30
18	136.5	4.4	2843	4	US-09-538-092-1007
19	136.5	4.4	2973	2	US-08-821-355A-7
20	136.5	4.4	2973	2	US-09-003-687A-7
21	136.5	4.4	2973	3	US-09-136-605-7
22	136	4.4	1139	4	US-09-949-016-6556
23	136	4.4	1146	4	US-09-949-016-7567
24	130.5	4.2	440	4	US-09-949-016-9774
25	128	4.1	601	4	US-09-252-991A-27821
26	120	3.9	427	4	US-09-252-991A-22649
27	117	3.8	329	4	US-09-252-991A-19752

28	115.5	3.7	458	3	US-09-352-159-9	Sequence 9, Appli
29	115.5	3.7	458	3	US-09-352-168-9	Sequence 9, Appli
30	115.5	3.7	458	4	US-09-771-045B-9	Sequence 9, Appli
31	115.5	3.7	458	4	US-09-770-564A-9	Sequence 9, Appli
32	113	3.7	432	4	US-09-252-991A-30848	Sequence 30848, A
33	112	3.6	622	4	US-09-252-991A-32308	Sequence 32308, A
34	112	3.6	2388	4	US-09-695-795A-2	Sequence 2, Appli
35	110.5	3.6	1972	4	US-08-875-435B-3	Sequence 3, Appli
36	110	3.6	326	4	US-09-252-991A-30438	Sequence 30438, A
37	109.5	3.5	2101	1	US-08-456-390-4	Sequence 4, Appli
38	109.5	3.5	2101	1	US-08-470-950-4	Sequence 4, Appli
39	109.5	3.5	2101	1	US-08-467-781-4	Sequence 4, Appli
40	109.5	3.5	2101	1	US-08-195-487-4	Sequence 4, Appli
41	109.5	3.5	2101	2	US-08-483-924-4	Sequence 4, Appli
42	109.5	3.5	2101	3	US-09-452-294-1	Sequence 1, Appli
43	109.5	3.5	2101	5	PCT-US93-06160-4	Sequence 4, Appli
44	108.5	3.5	2733	4	US-09-949-016-11433	Sequence 11433, A
45	108	3.5	541	4	US-09-252-991A-17206	Sequence 17206, A

ALIGNMENTS

RESULT 1

US-09-270-767-43983  
; Sequence 43983, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:

; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 43983  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-43983

Query Match	9.6%	Score 295.5;	DB 4;	Length 270;
Best Local Similarity	31.7%	Pred. No. 2.7e-21;		
Matches	64;	Conservative	46;	Mismatches 81; Indels 11; Gaps 3;
Qy	9	MSGXNKKQFLLERLLDAVKQCQIRFXGRKEIASDSRVTCCLCAQFAVLQHLKRS	68	
Db	48	LAGSTFGQAQDIFRRLQESAHKISQKFSG-KELATERDESQVCLCESLEBELMSYGLRQS	106	
Qy	69	RGALATAAA-----IKQAAGFASKTETEPVFWYVYKVEVKNKHLQRFYSLRHIASD	119	
Db	107	AGTSSFSAAAFIONMQEMVSGNAGGGS-NNNDATFWFCQTHLTPPFRQRQYMDLQIWTN	165	
Qy	120	VGRGRAWLRCALNEHSRLRYLHMLLADRCRLSFYEDWSFVMDERSMPLPTMAAGLSI	179	
Db	166	VGRGAFIRATLNEKRLSHVLTWLSDEEQHLRFYTPWLSLLNDEAKKLPFIVDSLSDV	225	
Qy	180	LFAINDNKDLNGSKFAPTVS	201	
Db	226	LFALNVDTTLNAPRRSTPSVA	247	

RESULT 2

US-09-370-838-191  
; Sequence 191, Application US/09370838  
; Patent No. 6444425  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Roadoh  
; APPLICANT: Secrist, Heather

Mon Jun 13 13:27:10 2005

```

; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370.838
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285.323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 191
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapien
; ORGANISM: Homo sapien
US-09-854-133-191

Query Match      4.9%; Score 151; DB 4; Length 704;
Best Local Similarity 21.3%; Pred. No. 5.3e-06;
Matches 111; Conservative 71; Mismatches 190; Indels 150; Gaps 20;

QY 56 QFEAVLQHLKRSRGLATAAAIKQAAGFASKTETEPVFWYVYVKEVLNKHQLQRFYSLRH 115
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 QFFVMEHCLK--HGLKVKKSFIGNKSPFGLVLEKLCPEASDIATS-----VRNLPE 195

QY 116 IASDVGRGAWLRCAALNEHSLERYLHMLADCRISTEYEDWSFVMDERSMPLPTMAAG 175
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 196 LKTAVGRGAWLYLALMOKKLADYLKVLIDNKHLLSEFYEPALMMEEB--GMVIVGLLVG 254

QY 176 LNSILFAINIDNKDLNGQ-----SKFAPTVDL--LKESTQNTVSLKKESTQGVSSLP 227
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 255 LNVLDANLCKGEDLDSQGVDFSLYLDVQDLGGKEH-ERITDVLQKNY----- 306

QY 228 EITASSAVSILIKPEQETDPCSCPGMSVLMFNNAKSGRRKXKPTXSHLMRKMRTILG 287
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 307 -----VEELNRHLSC---TVGDLQTKIDGLEK-----TNSKLOEELSAADR 345

QY 288 TCLKRLHGGQRAQTTPAPLISCPPLKAPSGLTPMESEQQLMENKXFFVFERGVWVPE- 346
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 346 IC-----SLOEQOQLREQNELIRERSEKSEVEI 373

QY 347 -----AXCEKHRRXXRCG-XKRRXRVWKLIRKQAQGLGVAREATGREHLPLDQALG-- 397
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 374 TKQDTKVELETYKTRQGLDEMYSVMKQKBEKKVRLELEKEL-----ELQIGMK 424

QY 398 -SAGRAAQXLRHPLPCQWRGLLQPSRCPKPRGDRTRGPRSPGWSWTSVQCGSQLSRPR 456
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 425 TEMEIAMKLL-----EKD-THEKQD-----TLVALRQOLEEVK 456

QY 457 KSSQPVTSASVPESM-----TISELRQATVAMNRKDBEENRSLRNLDDGEMHSA 510
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 457 AINLQMFHKAQNAESSLQOKNEAITSFEGKTNQVMSMKQMEERLQHSERARQGAERSH 516

QY 511 ALRQEV-----DTLKRKV-AEQEERQGMK 533
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 517 KIQQELGGRIGALQQLSQLHEQCSLEKELKSEKQEQALQ 558

RESULT 4
US-07-741-940-7
; Sequence 7, Application US/07741940
; Patent No. 5352775
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,940
; FILING DATE: 19920109
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; US-07-741-940-7

Query Match 4.4%; Score 136.5; DB 1; Length 2842;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;

QY 156 DWSFVMDERSSMLPTMAAGLNSILFAINDKDLGQSKFAPTVSDLLKESTONVTSLL 215
DB 2099 DWKAIQE-----GANSIVSLH-----QAAAAACLRSQASSDSLSL- 2137
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSGCGMSVLMNNAK----- 262
DB 2138 -KSGISLGSPFH-----LTPDQEEKPFTSNKGRPRILKPGKSTLETCKIESSEK 2185
QY 263 --RSGR-----KKKPTXSHLMRMKMSRTLTCLKHLGGQRAQT 302
DB 2186 GIKGGKKVYKSLITGKVRNSSEISGQMKQPLQANMPSISRGTW-----IHI-PGVNNS 2239
QY 303 TPTAPLSICPPLKAPSGLTPTMSESEQLMENXFPVFERGVWVPEACEKHXRCGXKRR 362
DB 2240 SSTSPVSKGPPPLKTPASKSPSEQTATTS-----PRGA-KPSVKSE----- 2280
QY 363 XRVWKLIRKBAQGPLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCQWRGL 417
DB 2281 --LSPVARQTSQ--IGGSSKAPRS-----GSRDSTPSRPAQQLSRPIQSPGRNS 2327
QY 418 LQPSR---CPPRKPGERDRTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464
DB 2328 ISPGRNISPPNKLSQLPRTSSP--STASTKSSGSGKMSYSPGQMSQQNLTKQTGLSKN 2386
QY 465 SASVPESMTTSE-LRQATVAM-MNRKDEL-----EENRSLNLLDGENEHSAA 512
DB 2387 ASSIPRESASKGLNQNNNGNKGANKVELSRMSSTKSSGSEDSRSPVL---VRQSTFI 2443
QY 513 QREVD-TLKRKVAE 525
DB 2444 KEAPSPTLRRKLEE 2457

RESULT 5
US-08-289-548A-7
; Sequence 7, Application US/08289548A
; Patent No. 5648212
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY

```

```

; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/289,548A
; FILING DATE: 12-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.46943
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; US-08-289-548A-7

Query Match 4.4%; Score 136.5; DB 1; Length 2842;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;

QY 156 DWSFVMDERSSMLPTMAAGLNSILFAINDKDLGQSKFAPTVSDLLKESTONVTSLL 215
DB 2099 DWKAIQE-----GANSIVSLH-----QAAAAACLRSQASSDSLSL- 2137
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSGCGMSVLMNNAK----- 262
DB 2138 -KSGISLGSPFH-----LTPDQEEKPFTSNKGRPRILKPGKSTLETCKIESSEK 2185
QY 263 --RSGR-----KKKPTXSHLMRMKMSRTLTCLKHLGGQRAQT 302
DB 2186 GIKGGKKVYKSLITGKVRNSSEISGQMKQPLQANMPSISRGTW-----IHI-PGVNNS 2239
QY 303 TPTAPLSICPPLKAPSGLTPTMSESEQLMENXFPVFERGVWVPEACEKHXRCGXKRR 362
DB 2240 SSTSPVSKGPPPLKTPASKSPSEQTATTS-----PRGA-KPSVKSE----- 2280
QY 363 XRVWKLIRKBAQGPLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCQWRGL 417
DB 2281 --LSPVARQTSQ--IGGSSKAPRS-----GSRDSTPSRPAQQLSRPIQSPGRNS 2327
QY 418 LQPSR---CPPRKPGERDRTRGPRSPGWSVQCG--SOLSRPKSSEQPV----- 464
DB 2328 ISPGRNISPPNKLSQLPRTSSP--STASTKSSGSGKMSYSPGQMSQQNLTKQTGLSKN 2386

```

QY	465	SASVPESTMWISL-LRQATVAM-MNRKDEL	-----EENRSLRNLLDGEHEHSAAL	512
Db	2387	ASSIPRSESASKGLNQMNGANGANKVELSRMSSTKSSGSESDRSRPVL	-----VRQSTFI	2443
QY	513	ROEVD-TLKRKVAE	525	
Db	2444	KEAPSTLRRKLEE	2457	
RESULT 6				
US-08-452-654-7				
; Sequence 7, Application US/08452654				
; Patent No. 5691454				
; GENERAL INFORMATION:				
; APPLICANT: ALBERTSEN, HANS				
; APPLICANT: ANAND, RAKESH				
; APPLICANT: CARLSON, MARY				
; APPLICANT: GRODEN, JOANNA				
; APPLICANT: HEDGE, PHILIP J.				
; APPLICANT: JOSLYN, GEOFF				
; APPLICANT: KINZLER, KENNETH				
; APPLICANT: MARKHAM, ALEXANDER F.				
; APPLICANT: NAKAMURA, YUSUKE				
; APPLICANT: THLIVERIS, ANDREW				
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC				
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS				
; NUMBER OF SEQUENCES: 94				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Banner, Birch, McKie & Beckett				
; STREET: 1001 G Street, NW				
; CITY: Washington				
; STATE: D.C.				
; COUNTRY: USA				
; ZIP: 20001-4598				
; COMPUTER READABLE FORM:				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: PatentIn Release #1.0, Version #1.25				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/452,654				
; FILING DATE: 25-MAY-1995				
; CLASSIFICATION: 536				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: US 07/741,940				
; FILING DATE: 08-AUG-1991				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Kagan, Sarah A.				
; REGISTRATION NUMBER: 32,141				
; REFERENCE/DOCKET NUMBER: 1107.035574				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 202-508-9100				
; TELEFAX: 202-508-9299				
; INFORMATION FOR SEQ ID NO: 7:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 2842 amino acids				
; TYPE: amino acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: protein				
; ORIGINAL SOURCE:				
; ORGANISM: Homo sapiens				
; IMMEDIATE SOURCE:				
; CLONE: APC				
US-08-452-654-7				
Query Match 4.4%; Score 136.5; DB 1; Length 2842;				
Best Local Similarity 21.9%; Pred. No. 0.0014;				
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;				
QY	156	DWSFVMDERSMPLPWAAGLSILFAINTDKDLNGQSKFAFTVSDLLKESQNTVTSLL	215	



```

;
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
;
; US-08-449-731-7
;
;
; Query Match 4.4% Score 136.5; DB 4; Length 2842;
; Best Local Similarity 21.9%; Pred. No. 0.0014;
; Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;
;
; QY 156 DWSFVMDERSMMLPTMAAGLSILFANINDKDLNQSKFAPTVSLLKSTQNVTSL 215
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2099 DWKAIQE-----GANSIVSSLH-----QAAAAACLRSQASSDSILSL- 2137
;
; QY 216 KESTQGVSSLFREITASSAVSILKPKQETDPCLSCPGMSVLPNNAK----- 262
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2138 -KSGISLGSPFH-----LTPDQEEKPFTSNKGPRIKPGEXSTLETKKIESEK 2185
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 263 --RSGRR-----KEXKPTXSHLMRMKMSRTLTGCTCKEHLGGQRAQT 302
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2186 GIQGGKVKYKSLITGKVRNSEISGQMKQLQANMPISGRTM-----IHI-PGVNNS 2239
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 303 TPTAPLISCPPLKAPSGLTPTMSESQQLMENXFPVFERGVVWPEACEKHXRCXGKRR 362
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2240 SFTSPSVKKGPLKTPASKPSFGQTATTS-----PRGA-KPSVKSE----- 2280
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 363 XRVWKLIRBAQGPLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCQWRGL 417
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2281 --LSPVARQTSQ--IGSSKAPSKS-----GSRDSTPSRPAQQPLSRPIQSPGRNS 2327
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 418 LQPSR---CPPRKPGEDRTRGPRSPQSWSVQCG--SQLSRPKSSEQPVV----- 464
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2328 ISPGNGISPPNKLSQLPRTSSP-STASTKSSGSGKMSYTPSGQMSQQNLTKQTGLSKN 2386
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EEENRSLRNLLDGEHEHAAL 512
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2387 ASSIPRSASAKGLNQNNNGANGANKVELSRMSSTKSSGSEDSRSERPVL---VRQSTFI 2443
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; QY 513 RQEVD-TLKRKVAE 525
; | | | | | | | | | | | | | | | | | | | | | | | | | |
; Db 2444 KEAPSPPTLRKLEE 2457
; | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Qy 418 LQPSR---CPPKPGERDRTRGPRSPGWSVTSCQC--SQLSRPKSSEQPVT----- 464
Db 2329 ISPNRNGISPPNKLSQLPRTSSP-STASTKSGSGRWSVTSPGRQWSQQNLTKQTGLSKN 2387
Qy 465 SASVPSPSWTISE-LRQATVAM-MNRKDEL-----EENRSRLNLLDGEHEMSAAL 512
Db 2388 ASSIPRSESASGLNOMNNGANKKVELSRMSSTKSGSGESDRSRPVL---VROSTFI 2444
Qy 513 ROEVD-TLKRKVAE 525
Db 2445 KEAPSTLRRKLEE 2458

RESULT 9
US-08-289-548A-2
; Sequence 2, Application US/08289548A
; Patent No. 5648212
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH

```

RESULT 8  
US-07-741-940-2  
; Sequence 2, Application US/07741940  
; Patent No. 5352775  
; GENERAL INFORMATION:  
; APPLICANT: ALBERTSEN, HANS  
; APPLICANT: ANAND, RAKESH  
; APPLICANT: CARLSON, MARY  
; APPLICANT: GRODEN, JOANNA  
; APPLICANT: HEDGE, PHILIP J.  
; APPLICANT: JOSLYN, GEOFF  
; APPLICANT: KINZLER, KENNETH  
; APPLICANT: MARKHAM, ALEXANDER F.  
; APPLICANT: NAKAMURA, YUSUKE  
; APPLICANT: THLIVERIS, ANDREW  
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Banner, Birch, McKie & Beckett

APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner & Allegretti, LTD  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/289,548A  
FILING DATE: 12-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.46943  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-289-548A-2

Query Match 4.4%; Score 136.5; DB 1; Length 2843;  
Best Local Similarity 21.9%; Pred. No. 0.0014;  
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;  
QY 156 DWSFVMDERSMLPTMAAGLSILFAINDKDLNGQSKFAPTVSDLLKSTQNTVSI 215  
DB 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLSRQASSDSDSILSL- 2138  
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMPPNAK----- 262  
DB 2139 -KSGISLGSPPH-----LTPDQEKFTSNKGPRIKPGKEKSTLETTKIESSEK 2186  
QY 263 --RSGR-----KRXXPTXSHLMRMKMSRTLTGTLKRLHIGQGRAQT 302  
DB 2187 GIKGGKVKYKSLITGKVRNSNISQMKQPLQANMPSISRGRTM-----IHI-PGVRNSS 2240  
QY 303 TTPATLSTCPPLKAPSGLTPTMSESQLMENXFPVFERGVVWPEACEKHXRCGXKR 362  
DB 2241 SSTSVSKGKPLPKTPASKSPSEGGTATTS-----PRGA-KPSVKSE----- 2281  
QY 363 XRWVKLIRKACGGLGVAREATGREHLPLPDALGSAEG-----AAQXLRHPLPCQWRGL 417  
DB 2282 --LSPVARTSQ--TGGSKAPRS-----GSRDSTPSPRPAQQLSRPIQSPGRNS 2328  
QY 418 LQPSR---CPPKRPGERDTRGPRSGWTSVQCG--SOLSPRKSSBPVPT----- 464  
DB 2329 ISFGNGISPPNKLSQLPRTSSP-STASTKSSGSGKMSYTSPPGRQMSQONLTKQTGLSKN 2387  
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSIRNLIDGEMEHSAL 512  
DB 2388 ASSIPRSESASKGLNQMNNGNANKKVELSRMSTKSGSGESDRSRPVL-----VRQSTFI 2444

QY 513 RQEVDTLKRKVAE 525  
DB 2445 KEAPSTLRRKLEE 2458  
RESULT 10  
US-08-452-654-2  
Sequence 2, Application US/08452654  
Patent No. 5691454  
GENERAL INFORMATION:  
APPLICANT: ALBERTSEN, HANS  
APPLICANT: ANAND, RAKESH  
APPLICANT: CARLSON, MARY  
APPLICANT: GRODEN, JOANNA  
APPLICANT: HEDGE, PHILIP J.  
APPLICANT: JOSLYN, GEOFF  
APPLICANT: KINZLER, KENNETH  
APPLICANT: MARKHAM, ALEXANDER F.  
APPLICANT: NAKAMURA, YUSUKE  
APPLICANT: THLIVERIS, ANDREW  
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC  
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Banner, Birch, McKie & Beckett  
STREET: 1001 G Street, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001-4598  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,654  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/741,940  
FILING DATE: 08-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A.  
REGISTRATION NUMBER: 32,141  
REFERENCE/DOCKET NUMBER: 1107.035574  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2843 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-452-654-2

Query Match 4.4%; Score 136.5; DB 1; Length 2843;  
Best Local Similarity 21.9%; Pred. No. 0.0014;  
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;  
QY 156 DWSFVMDERSMLPTMAAGLSILFAINDKDLNGQSKFAPTVSDLLKSTQNTVSI 215  
DB 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLSRQASSDSDSILSL- 2138  
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMPPNAK----- 262  
DB 2139 -KSGISLGSPPH-----LTPDQEKFTSNKGPRIKPGKEKSTLETTKIESSEK 2186  
QY 263 --RSGR-----KRXXPTXSHLMRMKMSRTLTGTLKRLHIGQGRAQT 302  
DB 2187 GIKGGKVKYKSLITGKVRNSNISQMKQPLQANMPSISRGRTM-----IHI-PGVRNSS 2240



; TITLE OF INVENTION: MUTATIONS IN THE APC GENE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BANNER & WITCOFF, LTD.  
; STREET: 1001 G STREET, N.W.  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: US  
; ZIP: 20001  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/370,235A  
; FILING DATE: 01-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KAGAN, SARAH A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 01107.48688  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 508 9100  
; TELEFAX: 202 508 9299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-370-235A-2

Query Match 4.4%; Score 136.5; DB 2; Length 2843;  
Best Local Similarity 21.9%; Pred. No. 0.0014;  
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;  
QY 156 DWSFVMDERSMLPTMAAGLSILFPAINDKDLGQSKFAPTVDLLKESTQNTVLSL 215  
DB 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLSRQASDSDSILSL- 2138  
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMENAK----- 262  
DB 2139 -KSGISLGSPFH-----LTPDQEEKPFTSNKGPRILKPGKSTLETTKIESK 2186  
QY 263 --RSGR-----KRKXPTYSHLMRMKMSRTLTGCLKRHLGQGRQRT 302  
DB 2187 GIKGKKVYKSLITGVRNSEISGQMKQPLQANMPSISRGRTM-----IHI-PGVRNS 2240  
QY 303 TPTAPLISCPPLKAPSLGTPMSEQOLMENKFPVFERGVWVPEAXCEKHXRCGXKR 362  
DB 2241 SSTSPVSKGPPPLKTPASKSPSEGQTATTS-----PRGA-KPSVKSE----- 2281  
QY 363 XRVWKLIRKEAQGLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCWRL 417  
DB 2282 --LSPVARQTSQ--IGSSKAPSKS-----GSRDSTPSRPAQPLSRPIQSPGRNS 2328  
QY 418 LQPSR---CPPRKGERDTRGPRSPGWSVQCG--SOLSRPKSSEQPV- 464  
DB 2329 ISPGRNGISPPNKLSQLPRTSSP-STASTKSGSGKMSYTPGROMSQOQLTKQTLGSKN 2387  
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRLLDGMEHSAAL 512  
DB 2388 ASSIPRESASKGLNQMNNGANGANKKVELSRMSTKSSGSESDRSERPVL---VRQSTFI 2444  
QY 513 ROEVD-TLKRKVAE 525  
DB 2445 KEAPSPTLRRKLEE 2458

RESULT 14  
US-08-450-582-2  
; Sequence 2, Application US/08450582

; FILING DATE: 25-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/289,548  
; FILING DATE: 12-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/741,940  
; FILING DATE: 08-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kagan, Sarah A.  
; REGISTRATION NUMBER: 32,141  
; REFERENCE/DOCKET NUMBER: 1107.49964  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-508-9100  
; TELEFAX: 202-508-9299  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2843 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ANTI-SENSE: NO  
; US-08-452-655B-7

Query Match 4.4%; Score 136.5; DB 1; Length 2843;  
Best Local Similarity 21.9%; Pred. No. 0.0014;  
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;  
QY 156 DWSFVMDERSMLPTMAAGLSILFPAINDKDLGQSKFAPTVDLLKESTQNTVLSL 215  
DB 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLSRQASDSDSILSL- 2138  
QY 216 KESTQGVSSLFREITASSAVSILIKPEQETDPCLSCPGMSVLMENAK----- 262  
DB 2139 -KSGISLGSPFH-----LTPDQEEKPFTSNKGPRILKPGKSTLETTKIESK 2186  
QY 263 --RSGR-----KRKXPTYSHLMRMKMSRTLTGCLKRHLGQGRQRT 302  
DB 2187 GIKGKKVYKSLITGVRNSEISGQMKQPLQANMPSISRGRTM-----IHI-PGVRNS 2240  
QY 303 TPTAPLISCPPLKAPSLGTPMSEQOLMENKFPVFERGVWVPEAXCEKHXRCGXKR 362  
DB 2241 SSTSPVSKGPPPLKTPASKSPSEGQTATTS-----PRGA-KPSVKSE----- 2281  
QY 363 XRVWKLIRKEAQGLGVAREATGREHLPLDPAQLGSAEG-----AAQXLRHPLPCWRL 417  
DB 2282 --LSPVARQTSQ--IGSSKAPSKS-----GSRDSTPSRPAQPLSRPIQSPGRNS 2328  
QY 418 LQPSR---CPPRKGERDTRGPRSPGWSVQCG--SOLSRPKSSEQPV- 464  
DB 2329 ISPGRNGISPPNKLSQLPRTSSP-STASTKSGSGKMSYTPGROMSQOQLTKQTLGSKN 2387  
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRLLDGMEHSAAL 512  
DB 2388 ASSIPRESASKGLNQMNNGANGANKKVELSRMSTKSSGSESDRSERPVL---VRQSTFI 2444  
QY 513 ROEVD-TLKRKVAE 525  
DB 2445 KEAPSPTLRRKLEE 2458

RESULT 13  
US-08-370-235A-2  
; Sequence 2, Application US/08370235A  
; Patent No. 5910418  
; GENERAL INFORMATION:  
; APPLICANT: VOGELSTEIN, BERT  
; APPLICANT: KINZLER, KENNETH W.  
; APPLICANT: HILL, DAVID E.  
; APPLICANT: JOHNSON, KAREN A.  
; TITLE OF INVENTION: ANTIBODIES AND ASSAYS FOR DETERMINING

```
Patent No. 6114124
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,582
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/452,655
FILING DATE: 25-MAY-1995
APPLICATION NUMBER: US 08/289,548
FILING DATE: 12-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.49964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-450-582-2
Query Match 4.4%; Score 136.5; DB 3; Length 2843;
Best Local Similarity 21.9%; Pred. No. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;
Qy 156 DWSFVMEDESSMLPTWAAAGNSILFPAINTDKDNGSKFAPTVSDLLKESTQNTVSL 215
Db 2100 DWKAIQE-----GANSIVSLH-----QAAAAACLSRQSSDSLSL- 2138
Qy 216 KESTQGVSSILFREITASSAVSILKPEQTDPCILSCPCMVLMPNAK----- 262
Db 2139 -KSGISIGSPH-----LTPQEEKPFTSNKGPRLKPKGEKSTLETKKIESEK 2186
Qy 263 --RSGR-----KRKXPTXSHLMRMKRSRTLTGCLKRHLGQGRQRT 302
Db 2187 GIKGKVKYKSLITGKVRNSNISEGQMKQPLQANMPSISRGRTM-----IHI-PGVNSS 2240
Qy 303 TPTAPLISCPPLKAPGLTPMESEQQLMENKFPVFERGVWVPEAKCEKHXXRCGKRR 362
Db 2241 SSTSPVSKGPPPLKTPASKSPSEQQTATTS-----PRGA-KPSVKSE----- 2281
363 XRVWKLIRKEAQGPGLGVAREATGREHLPLPDAQLGSAEG-----AAQXLRHPLPCQWRGL 417
2282 --LSPVARQTSQ--IGSSKAPSR-----GSRDSTPSRPAQQLSRPIQSPGRNS 2328
418 LQPSR---CPPRKGERDRTRGPRSPGSWTSVQCG--SLSRPRKSSSEQPVV----- 464
2329 ISPGNGISPNKLSQLPRTTSSP-STASTKSSGGRKMSYTSPPGQMSQQNLTKQTGLSKN 2387
465 SASVPESMTISE-LRQATVAM-MNRKDEL-----BEENRSLRNLLDGENEHSAAAL 512
2388 ASSIPRSESASKGLNQMNNGNGANKVELSRMSTKSSGESDSRSERPVL---VRQSTFI 2444
513 RQEVDTLKRKVAE 525
2445 KEAPSPTLRKLEE 2458
RESULT 15
US-08-450-582-7
Sequence 7, Application US/08450582
Patent No. 6114124
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,582
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/452,655
FILING DATE: 25-MAY-1995
APPLICATION NUMBER: US 08/289,548
FILING DATE: 12-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.49964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
```

Mon Jun 13 13:27:10 2005

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
US-08-450-582-7

Query Match      4.4%; Score 136.5; DB 3; Length 2843;
Best Local Similarity 21.9%; Pred. NO. 0.0014;
Matches 95; Conservative 60; Mismatches 140; Indels 139; Gaps 23;

QY 156 DWSFVWDEERSMLTMAAGLSILFAINIDKDLNGSKFAPTVDLLKESTQNTVSL 215
Db 2100 DWKAIQE-----GANSIVSSLH-----QAAAAACLRSQASSDSLSL- 2138
QY 216 KESTQGVSLFREITASSAVSILIKPEQETDPCLSCPQMSVLMPNAK----- 262
Db 2139 -KSGISLGSPPH-----LTPDQEKPFSTNKGPRILKPGCKSTLETKKIESK 2186
QY 263 --RSGR-----KRXPTYSHLMRKMSRTLTGCLKRHLGQGRAQRT 302
Db 2187 GIKGKKVYKSLITGKVRNSNISEIGMKQPLQANMPSISGRITM-----IHI-PGVRNS 2240
QY 303 TPTAPLISCPPLKAPSGLTPMSEQQLMENXFPVFERGVVWVPEACEKHXRCXGKRR 362
Db 2241 SSTSPVKKGPPKLPASKSPSEGQTATTS-----PRGA-KPSVKSE----- 2281
QY 363 XRVWKLIRKEAQGLGVAREATGREHLPLDQQLGSAEG-----AAQXLRHPLPCQWRGL 417
Db 2282 --LSPVARTSQ--IGGSKAPSR-----GSRDSTPSRPAQPLSRPIQSPGRNS 2328
QY 418 LQPSR---CPPRXGERDTRGPRSPGWSVQCG--SOLSRPKSSQOPVT----- 464
Db 2329 ISFGRNGISPPNKLSQLPTSSP-STASTKSGSGGKMSYTSFCRQMSQQLTKQTLGSKN 2387
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRNLLDGEHSAAL 512
Db 2388 ASSIPRESASAKGLNQMNNGANGKVELSRMSSTKSGSGESDRSRPVL---VRQSTFI 2444
QY 513 RQEVDTLKRKVAE 525
Db 2445 KEAPSTLRRKLEE 2458
```

Search completed: June 7, 2005, 12:04:48  
Job time : 25.6321 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: June 7, 2005, 12:02:31 ; Search time 76.4207 Seconds  
(without alignments)  
2846.907 Million cell updates/sec  
Title: US-09-155-676B-2  
Perfect score: 3093  
Sequence: 1 XTGPXGXMGSGXNDXKRF.....FLWTFHVCEPINCFSLSLKKK 604  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1599520 seqs, 360203123 residues  
Total number of hits satisfying chosen parameters: 1599520  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/FCF\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11A\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	183	5.9	1019	15	US-10-276-774-2298 Sequence 2298, Ap
2	175.5	5.7	473	15	US-10-094-749-1936 Sequence 1936, Ap
3	175.5	5.7	473	16	US-10-408-765A-2491 Sequence 2491, Ap
4	160	5.2	2845	15	US-10-267-502-372 Sequence 372, App
5	151	4.9	600	15	US-10-015-115-73 Sequence 73, Appl
6	151	4.9	704	9	US-09-738-973-191 Sequence 191, App
7	151	4.9	704	9	US-09-854-133-191 Sequence 191, App
8	151	4.9	704	14	US-10-144-649A-191 Sequence 191, App
9	145	4.7	546	15	US-10-264-237-2780 Sequence 2780, Ap
10	137	4.4	286	15	US-10-276-774-1412 Sequence 1412, Ap
11	137	4.4	286	15	US-10-296-115-1340 Sequence 1340, Ap
12	137	4.4	286	15	US-10-296-115-1465 Sequence 1465, Ap

13	136.5	4.4	369	15	US-10-264-237-2045 Sequence 2045, Ap
14	136.5	4.4	2843	8	US-08-681-219-32 Sequence 32, Appl
15	136.5	4.4	2843	9	US-09-987-482-1 Sequence 1, Appl
16	136.5	4.4	2843	10	US-09-230-111C-30 Sequence 30, Appl
17	136.5	4.4	2843	14	US-10-092-138-30 Sequence 30, Appl
18	136.5	4.4	2843	15	US-10-392-113-21 Sequence 21, Appl
19	136.5	4.4	2843	16	US-10-408-765A-1970 Sequence 1970, Ap
20	136.5	4.4	2843	16	US-10-820-403-30 Sequence 370, App
21	136.5	4.4	2844	15	US-10-267-502-370 Sequence 14, Appl
22	135.5	4.4	468	14	US-10-147-026-14 Sequence 4, Appl
23	124	4.0	494	14	US-10-035-343-4 Sequence 3547, Ap
24	124	4.0	745	15	US-10-108-260A-3547 Sequence 3340, Ap
25	124	4.0	787	15	US-10-104-047-3340 Sequence 2, Appl
26	124	4.0	1253	15	US-10-334-143-2 Sequence 2, Appl
27	123.5	4.0	1274	16	US-10-437-963-126375 Sequence 81, Appl
28	123	4.0	827	15	US-10-322-696-81 Sequence 234, App
29	122	3.9	827	15	US-10-428-275-234 Sequence 810, App
30	121.5	3.9	790	15	US-10-225-066A-810 Sequence 2384, Ap
31	121.5	3.9	790	15	US-10-374-780A-2384 Sequence 33, Appl
32	121	3.9	666	15	US-10-168-067C-33 Sequence 1, Appl
33	121	3.9	1190	15	US-10-168-067C-1 Sequence 61, Appl
34	120.5	3.9	1608	11	US-09-964-956-61 Sequence 60, Appl
35	120.5	3.9	1788	11	US-09-964-956-60 Sequence 63, Appl
36	120.5	3.9	2641	11	US-09-964-956-63 Sequence 1425, Ap
37	120.5	3.9	2641	16	US-10-408-765A-1425 Sequence 27, Appl
38	120.5	3.9	2811	11	US-09-964-956-27 Sequence 25, Appl
39	120.5	3.9	2814	11	US-09-964-956-25 Sequence 45308, A
40	119.5	3.9	520	15	US-10-425-114-45308 Sequence 183619, A
41	118	3.8	862	16	US-10-437-963-183619 Sequence 5585, Ap
42	118	3.8	1246	15	US-10-369-493-6585 Sequence 226, App
43	117	3.8	303	15	US-10-428-275-236 Sequence 230, App
44	117	3.8	827	15	US-10-428-275-222
45	117	3.8	827	15	US-10-428-275-230

ALIGNMENTS

RESULT 1

US-10-276-774-2298  
; Sequence 2298, Application US/10276774  
; Publication No. US20040053245A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-030  
; CURRENT APPLICATION NUMBER: US/10/276,774  
; PRIOR FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 2700  
; SOFTWARE: Custom  
; SEQ ID NO 2298  
; LENGTH: 1019  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-276-774-2298

Query Match 5.9%; Score 183; DB 15; Length 1019;  
Best Local Similarity 22.2%; Pred. No. 1.3e+06;  
Matches 115; Conservative 75; Mismatches 196; Indels 132; Gaps 25;

Qy	21	LLERLLDAVKQCIQIRFXG-RKEITAS--DSRVTCLCAQFENVLQGLKESRGLATAAA	77
Db	10	ILENISLSVKKLSYFAACEDEIPAIRNHDVKLQRLCEHLDHALLYGLQ-----	58
Qy	78	IKQAAGFASKTETPEPVFVYVVKLVKNKHEIQRFVSLRHIAASDVGRGRAWLRCALNHSLE	137
Db	59	-----DLSSGIWLVLVHFTTREAIKQIEVLQHVAINLGRSRAWLYLAUNENSLE	107









```

RESULT 10
US-10-276-774-1412
; Sequence 1412, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700

```





Db 2282 --LSPVARQTSQ--ICGSSKAPRS-----GSRDSTPSRPAQQPLSRPIQSPGRNS 2328  
QY 418 LOPSE---CPBRKPGERDTRGPRSPGSGWTSVQCG--SOLSRPRKSSFPVT----- 464  
Db 2329 ISPGRNGISPPNKLSQLPRTSSP-STASTKSGSGKMSYTPGROMSOQNLTKQTGLSKN 2387  
QY 465 SASVPESMTISE-LRQATVAM-MNRKDEL-----EENRSLRNLDDGEMEHSAL 512  
Db 2388 ASSIPRESASAKGLNQMNNGCANKKVELSRMSSTKSSGSESDRSERPVL---VRQSTFI 2444  
QY 513 ROEVD-TLKRKVAE 525  
Db 2445 KEAPSTLRRKLEE 2458

Search completed: June 7, 2005, 12:25:31  
Job time : 78.4207 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 894.226 Seconds  
(without alignments)  
17417.120 Million cell updates/sec

Title: US-09-155-676B-3  
Perfect score: 2631  
Sequence: 1 ccctctcacagccagcc.....acogtcgacctcgagggggg 2631

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2553.8	97.1	2631	2	AAV03324
2	1851.2	70.4	4596	2	AAV03326
3	1850.2	70.3	4596	5	AAV79490
4	1850.2	70.3	4596	6	ABK51171
5	1850.2	70.3	4596	10	ADK5075
6	1850.2	70.3	4596	13	ADR25114
7	1783.4	67.8	4684	11	ACN44439
8	1771.6	67.3	3152	12	ADQ67410
9	1133.4	43.1	3156	2	AAV73917
10	1133.4	43.1	3156	2	AAV71603
11	1133.4	43.1	3156	2	AAV69285
12	977	37.1	2844	2	AAV87842
13	973.4	37.0	73967	11	ACN44438
14	754.2	28.7	1385	4	AAH99262
15	691.8	26.3	2760	5	AAV79491
16	672.8	25.6	2828	11	ACN44437
17	588.2	22.4	722	5	AAV79492
18	446.8	17.0	2070	12	ADP28582
19	349.4	13.3	2475	11	ADM03069
20	242.2	9.2	499	9	ACH15373

21	205.8	7.8	476	10	ADF79802
22	188.8	7.2	42566	11	ACN44436
C 23	88.4	3.4	455	5	ABV22061
C 24	88.4	3.4	455	5	ABV27896
C 25	88.4	3.4	476	5	ABV12334
C 26	88.4	3.4	504	5	ABV33478
C 27	88.4	3.4	504	5	ABV42401
C 28	73.8	2.8	463	5	ABV03165
C 29	70	2.7	1024	5	AAV79493
C 30	58.2	2.2	2000	8	ADA71938
C 31	53.2	2.0	1253	2	AAV03325
C 32	52.8	2.0	65	6	ABV52988
C 33	50.4	1.9	374	4	AAI84793
C 34	50	1.9	50	6	ABZ03850
C 35	50	1.9	50	6	ABZ02592
C 36	49	1.9	60	6	ABN46309
C 37	43.6	1.7	2000	8	ADA71938
C 38	42.6	1.6	1464	6	AD24766
C 39	42.4	1.6	825	11	ABD10419
C 40	42.4	1.6	1359	11	ABD10504
C 41	42.4	1.6	1746	11	ABD10381
C 42	42	1.6	5280	12	ADN89158
C 43	41.2	1.5	1097	10	ADJ92320
C 44	40.4	1.5	1089	8	ACA26986
C 45	40.4	1.5	1711	1	AAV60310

ALIGNMENTS

RESULT 1

AAV03324  
ID AAV03324 standard; cDNA; 2631 BP.

XX AC AAV03324;

XX DT 15-APR-1998 (first entry)

XX DB Clone 10 cDNA encoding NMP1, a TRAF2 binding protein.

XX KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;

XX KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;

XX KW intracellular signalling activity; acute hepatitis;

XX KW autoimmune-induced cell death; ss.

XX OS Homo sapiens.

XX XX WO9737016-A1.

XX PD 09-OCT-1997.

XX XX 01-APR-1997; 97WO-IL000117.

XX PR 02-APR-1996; 96IL-00117800.

XX PR 26-AUG-1996; 96IL-00119133.

XX PA (YEDA ) YEDA RES & DEV CO LTD.

XX FI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX XX WPI; 1997-503101/46.

XX PT DNA encoding tumour necrosis factor receptor-associated factor binding

XX PT molecule - used for modulation or mediation in cells of the activity of

XX PT NF-KB.

XX PS Claim 4; Fig 4; 127pp; English.

XX CC The present sequence encodes a protein designated NMP1, a TRAF2 binding

XX CC protein. This protein contains Ser/Thr protein kinase motifs. A full

XX CC length cDNA clone encoding NF-kappaB (NIK) was obtained by PCR using the

XX CC present clone, clone 10. The clone 10 protein is capable of binding to at

XX CC least amino acids 222-501 of TRAF2. The TRAF-2 binding proteins can be

CC used for modulation or mediation in cells of the activity of NF-kappaB or  
 CC any other intracellular signalling activity modulated or mediated by  
 CC TRAF2. TRAF-binding proteins are especially used for prevention or  
 CC treatment of pathological conditions associated with NF-kB induction,  
 CC e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the  
 CC beta langerhans cells or the pancreas that results in diabetes, the death  
 CC of cells in graft rejection, the death of oligodendrocytes in the brain  
 CC in multiple sclerosis, and AIDS-inhibited T cell suicide which causes  
 CC proliferation of the AIDS virus and hence the AIDS disease. The proteins  
 CC are also useful for screening of ligands capable of binding to a protein,  
 CC which are useful for modulating cellular activity modulated/mediated by  
 CC TRAF2

XX	Sequence	2631 BP;	596 A;	791 C;	704 G;	459 T;	0 U;	81 Other;	
QY	Query Match	97.1%;	Score	2553.8;	DB	2;	Length	2631;	
QY	Best Local Similarity	100.0%;	Pred.	No.	0;				
QY	Matches	2631;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
QY	1	CCCCCTCTCAGCCAGCCAGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGCGTGT	60						
DB	1	CCCCCTCTCAGCCAGCCAGCCATCCAGAGGGCTGAGGAAGAGCCCATCCACCGCGTGT	60						
QY	61	CTGACGCGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA	120						
DB	61	CTGACGCGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA	120						
QY	121	GGCCTTGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	180						
DB	121	GGCCTTGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	180						
QY	181	ACCAGACCTTCAATGCCAGCGAGAGAGCTTTCCGCCAGGGCCGCCAGGGCCCGGCCAG	240						
DB	181	ACCAGACCTTCAATGCCAGCGAGAGAGCTTTCCGCCAGGGCCGCCAGGGCCCGGCCAG	240						
QY	241	CTGAGGAGACAAAGGAGAGCCCTTAAGCTCAGCCTCTCTCCACAGAGCCCGCCAG	300						
DB	241	CTGAGGAGACAAAGGAGAGCCCTTAAGCTCAGCCTCTCTCCACAGAGCCCGCCAG	300						
QY	301	AGCCAAAGAGTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGTGGGAACCTT	360						
DB	301	AGCCAAAGAGTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGTGGGAACCTT	360						
QY	361	TACCTCTGCTCCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420						
DB	361	TACCTCTGCTCCTCCCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420						
QY	421	CAACCGTCCCGAGCAGGAATCTCAGAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT	480						
DB	421	CAACCGTCCCGAGCAGGAATCTCAGAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT	480						
QY	481	CCAGCCATTTTCTCTGGAGGAGCAGGACAAATCTCTGTCCTCAGCATCGACAGCC	540						
DB	481	CCAGCCATTTTCTCTGGAGGAGCAGGACAAATCTCTGTCCTCAGCATCGACAGCC	540						
QY	541	TCTCCCTGTGCGATGACAGTGAGAAGAAACCCATCAAGGCTCTCAAGCTCCGCGGACA	600						
DB	541	TCTCCCTGTGCGATGACAGTGAGAAGAAACCCATCAAGGCTCTCAAGCTCCGCGGACA	600						
QY	601	CCCTGAGCTCAGGCGTACATCTCTGGAGAGCAGGCGGAGGCTCGAAGCTCCAGCTGGA	660						
DB	601	CCCTGAGCTCAGGCGTACATCTCTGGAGAGCAGGCGGAGGCTCGAAGCTCCAGCTGGA	660						
QY	661	ACATGGTGTGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG	720						
DB	661	ACATGGTGTGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG	720						
QY	721	TCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCTCGGAGATTCCACCGGGTCAAAG	780						
DB	721	TCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCTCGGAGATTCCACCGGGTCAAAG	780						
QY	781	TGGAGACATCGCCACTGGCATCAGAGCCAGATCCCGAGCTGCGAGCTTCAGCTTGTGTCA	840						
DB	781	TGGAGACATCGCCACTGGCATCAGAGCCAGATCCCGAGCTGCGAGCTTCAGCTTGTGTCA	840						

DB	781	TGGAGACATCGCCACTGGCATCAGAGCCAGATCCCGAGCTTCAGCTTTCAGCTTGGTCA	840
QY	841	CCAAAGCGGCGAGCCTGTTGCTACGATGAGGTGCGAGCTCGGCGATCGACCTGC	900
DB	841	CCAAAGCGGCGAGCCTGTTGCTACGATGAGGTGCGAGCTCGGCGATCGACCTGC	900
QY	901	AGTGCACACTGGCCCTTGATGGAGCTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGC	960
DB	901	AGTGCACACTGGCCCTTGATGGAGCTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGC	960
QY	961	TGAGAAACAGCCCTTAACCTCCCTCCACCGCGGCTCCACACTGCGGAAAGAGCCT	1020
DB	961	TGAGAAACAGCCCTTAACCTCCCTCCACCGCGGCTCCACACTGCGGAAAGAGCCT	1020
QY	1021	TCCTGCTCGTGCACGATGCTGCCCTGAAACACAGGCTCAGCCGTTTCCAGGGGATYT	1080
DB	1021	TCCTGCTCGTGCACGATGCTGCCCTGAAACACAGGCTCAGCCGTTTCCAGGGGATYT	1080
QY	1081	NCCAGCCCCCGGCTCARCAGNTGGGAACCCAGGGCTTCGNCAGCAGCAGGTTNGGGG	1140
DB	1081	NCCAGCCCCCGGCTCARCAGNTGGGAACCCAGGGCTTCGNCAGCAGCAGGTTNGGGG	1140
QY	1141	CAAGCAGAGATGCTCCAGGATTTTACANCTTGAGCCCTGCCANCCCTGCTGAADA	1200
DB	1141	CAAGCAGAGATGCTCCAGGATTTTACANCTTGAGCCCTGCCANCCCTGCTGAADA	1200
QY	1201	AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGGAGTTNNAACCTTGGGAA	1260
DB	1201	AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGGAGTTNNAACCTTGGGAA	1260
QY	1261	ACAAAACAGGATCTTTTCTGCCCCCTTCCAGTTCGAGTTGGCTGACCCCTTGG	1320
DB	1261	ACAAAACAGGATCTTTTCTGCCCCCTTCCAGTTCGAGTTGGCTGACCCCTTGG	1320
QY	1321	ANTCAGTGACCATTTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG	1380
DB	1321	ANTCAGTGACCATTTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG	1380
QY	1381	GTGGCGAGCCCTTGGCCCTTCAACCTNCCAGGCTGCTGAGAGTGTCAAGTGTGTA	1440
DB	1381	GTGGCGAGCCCTTGGCCCTTCAACCTNCCAGGCTGCTGAGAGTGTCAAGTGTGTA	1440
QY	1441	AGGNCACCAANCTCAGNTTTCAGTGAGAACAGGAGTTCAGAGCTTTCAGCCTGGGT	1500
DB	1441	AGGNCACCAANCTCAGNTTTCAGTGAGAACAGGAGTTCAGAGCTTTCAGCCTGGGT	1500
QY	1501	GTTTAAANNNGGGGCGCTCTNAAACCCCTTGCCTNGGCTNCAACCTNGGCCAGCTC	1560
DB	1501	GTTTAAANNNGGGGCGCTCTNAAACCCCTTGCCTNGGCTNCAACCTNGGCCAGCTC	1560
QY	1561	CTTTTGGGTGTAGGGGAAAGATGCTGAGCCCTGGGAGGCTTCCCTGGGTAGATACAC	1620
DB	1561	CTTTTGGGTGTAGGGGAAAGATGCTGAGCCCTGGGAGGCTTCCCTGGGTAGATACAC	1620
QY	1621	CACACTTTTTCAGTTGTTGCAACAGAGTCTGAGTTGACCTTCTGTTTTCAGCAAGG	1680
DB	1621	CACACTTTTTCAGTTGTTGCAACAGAGTCTGAGTTGACCTTCTGTTTTCAGCAAGG	1680
QY	1681	AAAGAGAGTGTAAAGTGAAGTGTCTGATNCCCAAGATGATGCCCCCTTGTGCTGT	1740
DB	1681	AAAGAGAGTGTAAAGTGAAGTGTCTGATNCCCAAGATGATGCCCCCTTGTGCTGT	1740
QY	1741	GCTACCACTTTCAGGAGAGAGGCGCCGAGCCCTTTCAGGCGCCAGCAGCTGCCCCAG	1800
DB	1741	GCTACCACTTTCAGGAGAGAGGCGCCGAGCCCTTTCAGGCGCCAGCAGCTGCCCCAG	1800
QY	1801	ACTCGCTGCACTCAGTTTCCCTCATCTGTAAAGTGAAGGTTGATGAGGATATGCTGTA	1860
DB	1801	ACTCGCTGCACTCAGTTTCCCTCATCTGTAAAGTGAAGGTTGATGAGGATATGCTGTA	1860
QY	1861	CAGGAACAGTCTGTGTGATGGAATGATCAGTCTTAAAGGAGGAGAGAGAGAGAG	1920
DB	1861	CAGGAACAGTCTGTGTGATGGAATGATCAGTCTTAAAGGAGGAGAGAGAGAGAG	1920





Db 2519 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAAATAGAAATTATTCTCTCAACAGCCTGT 2578  
 Qy 481 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCTCAGCATCGACAGCC 540  
 Db 2579 CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCTCAGCATCGACAGCC 2638  
 Qy 541 TCTCCTGTGCGATGACATGAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGAC 600  
 Db 2639 TCTCCTGTGCGATGACATGAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGAC 2698  
 Qy 601 CCTCAGCTCAGCGTACACTCTCGAGCAGCGAGCCAGGCTCGAAGCTCCAGCTGGA 660  
 Db 2699 CCTCAGCTCAGCGTACACTCTCGAGCAGCGAGCCAGGCTCGAAGCTCCAGCTGGA 2758  
 Qy 661 ACATGGTGTGCTCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720  
 Db 2759 ACATGGTGTGCTCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2818  
 Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCGGAGTTCCACCGGTTCAAAG 780  
 Db 2819 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCGGAGTTCCACCGGTTCAAAG 2878  
 Qy 781 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGGTCA 840  
 Db 2879 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGGTCA 2938  
 Qy 841 CCAAGACCGGAGCCTGTTTCGTAGCATGAGAGTGCAGACTCGGGCATCGACCTGC 900  
 Db 2939 CCAAGACCGGAGCCTGTTTCGTAGCATGAGAGTGCAGACTCGGGCATCGACCTGC 2998  
 Qy 901 AGTGACACATGCGCCCTGATGCGAGCTTCGCTGAGCTGGAGGTCAAGCATGCGCCAGC 960  
 Db 2999 AGTGACACATGCGCCCTGATGCGAGCTTCGCTGAGCTGGAGGTCAAGCATGCGCCAGC 3058  
 Qy 961 TGGAGAACCGCCCTAACCTTGCCTTCCACCGCGCTCCACACTGCGGAAAGCAGCT 1020  
 Db 3059 TGGAGAACCGCCCTAACCTTGCCTTCCACCGCGCTCCACACTGCGGAAAGCAGCT 3118  
 Qy 1021 TCTGTCTGGTGACATGCTGCCCTGAAACACAGGCTCAGCGTTCCTCCAGGGATGTG 1080  
 Db 3119 TCTGTCTGGTGACATGCTGCCCTGAAACACAGGCTCAGCGTTCCTCCAGGGATGTG 3176  
 Qy 1081 NCCAGCCCCCGCTCARCAGNTGGGAAACAGGGCCTCGNACAGNAGCTGNGGGG 1140  
 Db 3177 GCCAGCCCCCGCTCA--CAGTGGGAAACAGGGCCTCG-----CAGCAGCAAGGTGGG 3229  
 Qy 1141 CAAGCNAGATGCTCTCCAGGATTTCAANCTGAGCCNCTGCCANCCCTGCTCAADA 1200  
 Db 3230 GCAAGCAGATGCTCTCCAGGATTTCAACCTGAGCCCTGCCCCAC-----CCTGCTGA 3283  
 Qy 1201 AAACATNCCGACAGTGAAGACAGAGGAGGATGGNACGAGTTNNAACCTTGGGGAA 1260  
 Db 3284 AAACATNCCGACAGTGAAGACAGAGGAGGATGGC-----AGGAGTTACCTGGGGAA 3339  
 Qy 1261 AAAAAACAGGGATCTTTTCTGCCCCCTGCTCAGTNCAGTTGGCTGNAACCCGCTGG 1320  
 Db 3340 AAAAAACAGGGATCTTTTCTGCCCCCTGCTCAGT--CGAGTTGGGCTGA--CCCCGTTG 3395  
 Qy 1321 ANTCAGTGACATTTGTTGGCAGNACAGGGAGAGCAGCTTCCAGCTGGGTGAGAGGG 1380  
 Db 3396 GATCAGTGACATTTGTTGGCAGA--CAGGGAGAGCAGCTTCCAGCTGGGTGAGAGGG 3454  
 Qy 1381 GTGGGCGAGCCCTTCGGCCCTCACCTNCCAGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440  
 Db 3455 GTGGGCGAGCCCTTCGGCCCTCACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGTA 3512  
 Qy 1441 AGGNCNCAANCTCAAGNTTCAGTGAGAACCAAGTNCAGAGGATGATGCCGCCGNTA 1500  
 Db 3513 AGGNCNCAANCTCAAGNTTCAAGTTCAGTGCAGAACCAAGTNCAGAGGATGATGCCGCCGNTA 3568  
 Qy 1501 GGTTAANNNGGGGCCCTCTNAAACCCCTTGTCTGCTNNGCTNCACTNCGCCAGCTCANCCC 1560  
 Db 3569 GTTAA-----GGGGGCCCTCTNAAACCCCTTGTCTGCTNNGCTNCACTNCGCCAGCTCA--CCC 3620

Qy 1561 CTTTTCGGTCTAGGGAAAAGATGCTGACCTCTGGAAAGGCTTWCCTCTGTAGAAATACAC 1620  
 Db 3621 CTTTTCGGTCTAGGGAAAAGATGCTGACCTCTGGAAAGGCT--CCTCTGTAGAAATACAC 3679  
 Qy 1621 CACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 1680  
 Db 3680 CACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 3739  
 Qy 1681 AAAGAGGTGTGTAAGTGAAGTGTCTCAGTNCCTCCAGACATGTCCTCTTGTGCTGCTG 1740  
 Db 3740 AAAGAGGTGTGTAAGTGAAGTGTCTCAGT--CCCAGACATGTCCTCTTGTGCTGCTG 3798  
 Qy 1741 GCTTACCACTCTTCCAGAGCAGCGCCCGAGCCCTTTCAGGCCACGACACTGCCCCAG 1800  
 Db 3799 GCTTACCACTCTTCCAGAGCAGCGCCCGAGCCCTTTCAGGCCACGACACTGCCCCAG 3858  
 Qy 1801 ACTGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGTA 1860  
 Db 3859 ACTGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGTA 3918  
 Qy 1861 CAGGAACAGTCTGTGGATGACATGATCAGTGTCTNAGGNAAGCAGCAGAGAGAGAGCY 1920  
 Db 3919 CAGGAACAGTCTGTGGATGACATGATCAGTGTCTNAGGNAAGCAGCAGAGAGAGAGCY 3975  
 Qy 1921 TCCGCGCCCCCAGNCCCACCTNATCAGTGNCCAGCCTGTCTGTTNCCCAGNAGAC 1980  
 Db 3976 TCCGCGCCCCCAGNCCCACCTNATCAGTGT---CCAGCGTGTGTTTCCCCAGAGCAGC- 4031  
 Qy 1981 GCTNCAAGCATCAGCTGACATNCACTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
 Db 4032 -----TCAGCATCAGCTGACACTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4082  
 Qy 2041 CCGNACGCACTTTCGACNTCTGATGNACCTTCAAGCAGCTTTCATGCTGCTGCTGCTG 2100  
 Db 4083 CCG--ACGCACTTTCGACT--CTGATGACCTTCAAGCAGCTTTCATGCTGCTGCTGCTG 4139  
 Qy 2101 GCAGGNCAGGNCAGGNCAGTGCATCTGTAGNAGCATANGCAAGCCAGAGATGG 2160  
 Db 4140 GGGCAGGCGCAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4189  
 Qy 2161 GGTNAGGAGNACAGCTTTCGAGCTGTCANCATGATGCTGCTGCTGCTGCTGCTGCTG 2220  
 Db 4190 GGTGAA--GGGACACAGCTTTCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4244  
 Qy 2221 NCCAGNATTTCTTAAGAAATAGCAGCCCTTTCGCTGCTGCTGCTGCTGCTGCTGCTG 2280  
 Db 4245 --CCAGATTTCTTAAGAAATAGCAGCCCTTTCGCTGCTGCTGCTGCTGCTGCTGCTG 4300  
 Qy 2281 CCAGGAGGAGTCTCAGGACTCAGTACATTAATTAATTAATTAATTAATTAATTAATTA 2340  
 Db 4301 CCAGGAGGAGTCTCAGGACTCAGTACATTAATTAATTAATTAATTAATTAATTAATTA 4358  
 Qy 2341 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
 Db 4359 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4418  
 Qy 2401 TCCCACTCTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460  
 Db 4419 TCCCACTCTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4477  
 Qy 2461 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520  
 Db 4478 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4537  
 Qy 2521 GGTGTTGAGAGACTCTTTTGTAAAGCAATAAGTTTGGGTTGATGACAAATGTTAAAAA 2579  
 Db 4538 GGTGTTGAGAGACTCTTTTGTAAAGCAATAAGTTTGGGTTGATGACAAATGTTAAAAA 4596

XX AAS79490;  
AC  
XX 13-FEB-2002 (first entry)  
DT  
XX  
XX DNA encoding novel human diagnostic protein #15294.  
DE  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
KW  
XX Homo sapiens.  
OS  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
PR  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX WPI; 2001-639362/73.  
DR  
XX P-PSDB; ABG15303.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 15294; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: the sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;  
SQ

Query Match 70.3%; Score 1850.2; DB 5; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;  
QY 1 CCCCTCTCACAGCCAGCCATCAAGAGGGCTGAGAGAGAGCCATCCACCGCGTGT 60  
DB 2100 CCCCTCTCACAGCCAGCCATCAAGAGGGCTGAGAGAGAGCCATCCACCGCGTGT 2159  
QY 61 CTGAGCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAGTGGAGGTCTGAAGA 120  
DB 2160 CTGAGCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAGTGGAGGTCTGAAGA 2219  
QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATPACC 180

DB 2220 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATPACC 2279  
QY 181 ACCAGACCTCCATGCCCCAGCCGAGAGAGCTTTTCGCAAGGGGCCCGAGGGCCCCGGCCAG 240  
DB 2280 ACCAGACCTCCATGCCCCAGCCGAGAGAGCTTTTCGCAAGGGGCCCGAGGGCCCCGGCCAG 2339  
QY 241 CTGAGGAGACAAAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGAGCCCCCAG 300  
DB 2340 CTGAGGAGACAAAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGAGCCCCCAG 2399  
QY 301 AGCCAAACAAGTCT 360  
DB 2400 AGCCAAACAAGTCT 2459  
QY 361 TACCTCTGTCT 420  
DB 2460 TACCTCTGTCT 2519  
QY 421 CAACCGTCCCGAGAGAGAACTGCGAGAGCTGGAATAGAAATATTTCTCTCAACAGCCTGT 480  
DB 2520 CAACCGTCCCGAGAGAGAACTGCGAGAGCTGGAATAGAAATATTTCTCTCAACAGCCTGT 2579  
QY 481 CCAGAGCATTTCT 540  
DB 2580 CCAGAGCATTTCT 2639  
QY 541 TCTCCCTGTCT 600  
DB 2640 TCTCCCTGTCT 2699  
QY 601 CCTGAGCTCAGCGGTACATCTCTGAGAGAGCCAGCGGAGGCTCGAAGTCTCAGCTTGA 660  
DB 2700 CCTGAGCTCAGCGGTACATCTCTGAGAGAGCCAGGCGGAGGCTCGAAGTCTCAGCTTGA 2759  
QY 661 ACATGGTCTGGCCCGGGGGGGCCACCCAGACATCCCAAGGCTCTCTCAAGTCTGTTGAAG 720  
DB 2760 ACATGGTCTGGCCCGGGGGGGCCACCCAGACATCCCAAGGCTCTCTCAAGTCTGTTGAAG 2819  
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGCTCAAG 780  
DB 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGCTCAAG 2879  
QY 781 TGGGAGACATCGCCATCGGATCAGAGCCAGATCCAGCTGACGCTTCAAGTCTGTTGA 840  
DB 2880 TGGGAGACATCGCCATCGGATCAGAGCCAGATCCAGCTGACGCTTCAAGTCTGTTGA 2939  
QY 841 CCAAAGAGGGGAGCTGTTGCTACGATGAGAGTCCAGACTCGGGCATCAACCTGC 900  
DB 2940 CCAAAGAGGGGAGCTGTTGCTACGATGAGAGTCCAGACTCGGGCATCAACCTGC 2999  
QY 901 AGTGACACTGGCCCTGTATGGGAGCTTCGCTGGAGTGGAGGTCAAGCATGCCAGC 960  
DB 3000 AGTGACACTGGCCCTGTATGGGAGCTTCGCTGGAGTGGAGGTCAAGCATGCCAGC 3059  
QY 961 TGGAGAACAGCGCCCTTAACCCCTCGCCCTCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020  
DB 3060 TGGAGAACAGCGCCCTTAACCCCTCGCCCTCACCGCGGCTCCACACTGCGGAAAGCAGCCT 3119  
QY 1021 TCCTGTCTGTGACAGATGTCCTGTGAAACACAGAGCTCAGCGCTTCCAGGGGATVTG 1080  
DB 3120 TCCTGTCTGTGACAGATGTCCTGTGAAACACAGAGCTCAGCGCTTCCAGGGGATVTG 3177  
QY 1081 NCCAGCCCCCGGCTCAGAGTGGGAAACAGAGGCTTCGACGACGACGACGACGACGACGACGAC 1140  
DB 3178 GCAGAGCCCCCGGCTCA--CAGTGGGAAACAGAGGCTTCG-----CAGCAGCAAGGTGGGG 3230  
QY 1141 CAAGCNAGATGCTCTCCAGGATTTTCAACCTGAGCCCTTGCCTCCACCTCTCTGTA 1200  
DB 3231 GCAGAGCATGCTCTCCAGGATTTTCAACCTGAGCCCTTGCCTCTCTGTA 3284  
QY 1201 AAACAYTNCGCCACGTGAAGAGAGACAGAGGAGATCGNCAGGAGTNNACCTTGGGGAA 1260  
DB 3285 AAAACATCCGCCACGTGAAGAGAGACAGAGGAGTGGC---AGGAGTTACCTGGGGAA 3340



Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCCTCTCACAGCCAGCCCATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60  
Db 2100 CCCCTCTCACAGCCAGCCCATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 2159  
Qy 61 CTGAGCGGAGCTGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
Db 2160 CTGAGCGGAGCTGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219  
Qy 121 GCCCTTGGAGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC 180  
Db 2220 GCCCTTGGAGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC 2279  
Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCAGGGCCCGGCCAG 240  
Db 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCAGGGCCCGGCCAG 2339  
Qy 241 CTGAGGAGACAAGGCGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGCCCGCAG 300  
Db 2340 CTGAGGAGACAAGGCGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGCCCGCAG 2399  
Qy 301 AGCCAAACAAGTCTCCCTTGTACTTTGACAGAGGAGGAGTCTGGGATGTGGAAACCTT 360  
Db 2400 AGCCAAACAAGTCTCCCTTGTACTTTGACAGAGGAGGAGTCTGGGATGTGGAAACCTT 2459  
Qy 361 TACCTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAG 420  
Db 2460 TACCTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAG 2519  
Qy 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATATCTCAACAGCCTGT 480  
Db 2520 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATATCTCAACAGCCTGT 2579  
Qy 481 CCGAGCCATTTTCTGGAGAGCAGGAGCAAAATCTCTGTGCTCAGCATCGACAGCC 540  
Db 2580 CCGAGCCATTTTCTGGAGAGCAGGAGCAAAATCTCTGTGCTCAGCATCGACAGCC 2639  
Qy 541 TCTCCCTGTGCGATGACAGTGAAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600  
Db 2640 TCTCCCTGTGCGATGACAGTGAAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699  
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGCGAGCTCGAAGCTTCAAGCTGGA 660  
Db 2700 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGCGAGCTCGAAGCTTCAAGCTGGA 2759  
Qy 661 ACATGCTGCTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGCTGTAAG 720  
Db 2760 ACATGCTGCTGGCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGCTGTAAG 2819  
Qy 721 TCCAAATACAGTCTCTTAATGCTGAACACCTGACATCCGGGAGTTCACCGGTCAAG 780  
Db 2820 TCCAAATACAGTCTCTTAATGCTGAACACCTGACATCCGGGAGTTCACCGGTCAAG 2879  
Qy 781 TGGAGACATCGCCACCTGGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGCTTGTGCA 840  
Db 2880 TGGAGACATCGCCACCTGGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGCTTGTGCA 2939  
Qy 841 CCAAAGACGGGACGCTTTCGCTACGACATGAGAGTGGCAGATCGGGCATCGACCTGC 900  
Db 2940 CCAAAGACGGGACGCTTTCGCTACGACATGAGAGTGGCAGATCGGGCATCGACCTGC 2999  
Qy 901 AGTGCACACTGGCCCTGTAGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTGCACACTGGCCCTGTAGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 3059  
Qy 961 TGGAGAACAGGCCCTTAACCTGCCCTCCACCGCGGCTCCACACTGCGGAAAGAGCCTT 1020  
Db 3060 TGGAGAACAGGCCCTTAACCTGCCCTCCACCGCGGCTCCACACTGCGGAAAGAGCCTT 3119  
Qy 1021 TCCTGCTCGGTGACAGTGTGCCCTGAAACACAGGCTCAGCCGTTCCAGGGGATYTG 1080  
Db 4141 GGGCAGGGCAGGCAGTGCACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190

Db 3120 TCCTGCTCGGTGACAGTGTGCTCCCTGAAACAACAGGCTCAGCCGTTCCACAGGGGAT-- 3177  
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCGCTTCGNACAGVAGCNAAAGTNGGGGG 1140  
Db 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACACAGGCGCTCG-----CAGCAGCAAGGTGGGG 3230  
Qy 1141 CAAGCNAGATGCTCTCCAGGATTTTCAACNCTGAGCCGNTGCCCNCCCTGCTGAADA 1200  
Db 3231 GCAAGCAGAAATGCTCTCCAGGATTTTCAACCTGAGCCCTGCCCCAC-----CCTGCTGA 3284  
Qy 1201 AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGGAGTTNNACCTTGGGGAA 1260  
Db 3285 AAACACATCCGCCACGTAAGAGACAGAGGAGATGGC-----AGGATTTACCTTGGGGAA 3340  
Qy 1261 AAAAAACAGGATCTTTTNTTCTGCCCTGTCTCCAGTNCAGTGTGGCCTGNACCCGCTTGG 1320  
Db 3341 AAAAAACAGGATCTTTT--TTCTGCCCTGTCTCCAGT--CGAGTTGGCCTGA--CCCGCTTG 3396  
Qy 1321 ANTCAGTGACATTTGTTGGCAGANCAAGGGGAGAGAGCTTTCAGCCCTGGGTCAAGAGGG 1380  
Db 3397 GATCAGTGACCATTTGTTGGCAGA-CAGGGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGGG 3455  
Qy 1381 GTGGCGAGCCCTTCGGCCCTCACCCCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440  
Db 3456 GTGGCGAGCCCTTCGGCCCTCACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGTA 3513  
Qy 1441 AGGNCCTCAAACTCAGGNTTCAGTGCAGAACACAGGTNACAGCAGTATGCCCCGCCGNTA 1500  
Db 3514 AGGCCCAAACTC--AGGTTCAGTGCAGAACACAGGT--CAGCAGTATGCCCCCGCTAG 3569  
Qy 1501 GGTAAANNNGGGGCGCTCTNAAACCCCTTGCTGCTGCTGCTNCACTNNGGCCAGCTCANCCC 1560  
Db 3570 GTTAA-----GGGGCGCCCTCTAAACCCCTTGCTGGCCTCA---CCTGGCCAGCTCA--CCC 3621  
Qy 1561 CTTTGGGCTAGGGNAAGAAATGCTGACCTCTGGGAAGGCTTCCTGCTGTAGATACAC 1620  
Db 3622 CTTTGGGCTAGGGNAAGAAATGCTGACCTCTGGGAAGGCT--CCTTGGTAGAATACAC 3680  
Qy 1621 CACACTTTTTCAGGTTGTGCAACACACAGGCTCTGAGTTGACCTCTGTGTTACGCCAAGACC 1680  
Db 3681 CACACTTTTTCAGGTTGTGCAACACACAGGCTCTGAGTTGACCTCTGTGTTACGCCAAGACC 3740  
Qy 1681 AAAGAGGTGTGAAGTGAAGTGTCTCAGTNCNCCAGACATGTGCCCCCTTTCTGCTGTG 1740  
Db 3741 AAAGAGGTGTGAAGTGAAGTGTCTCAGT--CCCCAGACATGTGCCCCCTTTCTGCTGTG 3799  
Qy 1741 GCTACCACTCTTCCCCAGAGCAGCGCCCGAGCCCTTTCAGGCCCAGCACTGCCCCAG 1800  
Db 3800 GCTACCACTCTTCCCCAGAGCAGCGCCCGAGCCCTTTCAGGCCCAGCACTGCCCCAG 3859  
Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGACGATATGCTGTA 1860  
Db 3860 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGACGATATGCTGTA 3919  
Qy 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTNAAAGNAAAGCAGCAGAGAGACGY 1920  
Db 3920 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTGTA--GGAAAGCAGCAGAGAGAGC- 3976  
Qy 1921 TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTGTTNCCCGAGNAGACA 1980  
Db 3977 TCCGGCCCCCAGNCCCACTNATCAGTGT--CCAGCGTGTGTTTCCCGAGAGCAGC- 4032  
Qy 1981 GCTNAGNCACTCANCACTNCACTNCCCTNGCCCTGCCCCCTNGGCCANAGAGGTA 2040  
Db 4033 -----TCAGCATCACTGACACTCACTCACCCTGCCCCCTG-----CCAGAGGGTACTG 4083  
Qy 2041 CCGNACGGCACTTTTGCACNTCTGATGNACTCAAGACATTTTTCATGCTNGCCCTCTNNG 2100  
Db 4084 CCG-ACGGCACTTTTGGCACT--CTGATGACCTCAAGACACTTTTCATGCTGCTCTGGCA 4140  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGGNCATANGCAAGCCAGGAGATGG 2160  
Db 4141 GGGCAGGGCAGGCAGTGCACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190

```
QY 2161 GGTGNAAGGGANACAGTCTTGAGCTGTCCANCATGTCATGTGACTTNCCTCAAACTCTTN 2220
    |||||
Db 4191 GGTGAA--GGGACACAGTCTTGAGCTGTCCA-CATGCAATGTGACT-CCTCAAACTCTTT- 4245
    |||||
QY 2221 NCCAGNATTCTCTAAGAAATAGCAGNCCCTTTNCCCTTTGCCCAGCTTAGCTCTTCT 2280
    |||||
Db 4246 --CCAGATTCTCTAAGAAATAGCAGCCCC--TTCCCCATTGCCAGCTTAGCTCTTCT 4301
    |||||
QY 2281 CCCAGGGGAGCTTANCTCAGGACTCAGTAGCATTAATCAGCTGTGNATCTGAGGGG 2340
    |||||
Db 4302 CCCAGGGGAGCTA-CTCAGGACTCAGTAGCATTAATCAGCTGTG-AATCGTCAAGGGG 4359
    |||||
QY 2341 TGTCTGCTAGCTCAACCTCTCTGGGCGAGGGGACGCCGAGACTCCGTTGGGAGAACTCAT 2400
    |||||
Db 4360 TGTCTGCTAGCTCAACCTCTCTGGGCGAGGGGACGCCGAGACTCCGTTGGGAGAACTCAT 4419
    |||||
QY 2401 TCCACATCTTCCCAAGACAGCCTTTNFTCCAGCTGTCCACATTCAGTTCAGCTCTCCC 2460
    |||||
Db 4420 TCCACATCTTCCCAAGACAGCCTTT-GTCCAGCTGTCCACATTCAGTTCAGCTCTCCC 4478
    |||||
QY 2461 GGGGAGAGAGCCCGGCCCCCAGCACATTAAGAACTTCGAGCTTGGTACTGCAGAGTCTG 2520
    |||||
Db 4479 GGGGAGAGAGCCCGGCCCCCAGCACATTAAGAACTTCGAGCTTGGTACTGCAGAGTCTG 4538
    |||||
QY 2521 GGTCTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 2578
    |||||
Db 4539 GGTCTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 4596
    |||||
```

## RESULT 5

ADE85075

ID ADE85075 standard; DNA; 4596 BP.

XX AC ADE85075;

XX AC ADE85075;

XX DT 29-JAN-2004 (first entry)

XX DE Farnesyl transferase inhibitor modulated leukemia associated gene #294.

XX KW ss; cytostatic; farnesyl transferase inhibitor; gene expression;

XX KW quinolone; leukemia; cancer.

XX OS Homo sapiens.

XX PN WO2003038129-A2.

XX XX 08-MAY-2003.

XX PF 30-OCT-2002; 2002WO-US034784.

XX PR 30-OCT-2001; 2001US-0338997P.

XX PR 30-OCT-2001; 2001US-0340081P.

XX PR 30-OCT-2001; 2001US-0340938P.

XX PR 30-OCT-2001; 2001US-0341012P.

XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.

XX PI Raponi M;

XX DR WPI; 2003-513497/48.

XX PT Determining whether a patient will respond to treatment with a farnesyl transferase inhibitor, by analyzing the expression of gene that is differentially modulated in the presence of the inhibitor.

XX PS Disclosure; SEQ ID NO 294; 346pp; English.

XX CC The invention relates to a method of determining whether a patient will respond to treatment with a farnesyl transferase inhibitor (FTI), by analyzing the expression of gene that is differentially modulated in the presence of an FTI. The method is useful for determining whether a patient will respond to treatment with a FTI such as (B)-6-(amino(4-

CC chlorophenyl) (1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-methyl-2-(1H)quinoline, monitoring the therapy of a patient, treating a patient with leukemia with FTI if the analysis indicates that the patient will respond. This sequence corresponds to a gene whose expression may be modulated in the presence of FTI.

XX SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;

Query Match 70.3%; Score 1850.2; DB 10; Length 4596;

Best Local Similarity 90.5%; Pred. No. 0;

Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCAGGCCCATCCAAGAGGGGCTGAGAAAGAGCCATCCACCGCGTGT 60

Db 2100 CCCCTCTCACAGCCAGGCCCATCCAAGAGGGGCTGAGAAAGAGCCATCCACCGCGTGT 2159

QY 61 CTGACGCGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120

Db 2160 CTGACGCGGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2219

QY 121 GCCCTTGGAGGGGAGATATTAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC 180

Db 2220 GCCCTTGGAGGGGAGATATTAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC 2279

QY 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 240

Db 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGGCCCGGCCAG 2339

QY 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGCCCGCCAG 300

Db 2340 CTGAGGAGACAAACAGGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGCCCGCCAG 2399

QY 301 AGCCAAACAAGTCTCCCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATGTGGAAACCT 360

Db 2400 AGCCAAACAAGTCTCCCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATGTGGAAACCT 2459

QY 361 TACCTCTGTCTCCCTGGAGCCAGCCCTTGCAAGAAACCCAGCTCACAGAGCGGAAAG 420

Db 2460 TACCTCTGTCTCCCTGGAGCCAGCCCTTGCAAGAAACCCAGCTCACAGAGCGGAAAG 2519

QY 421 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 480

Db 2520 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGT 2579

QY 481 CCCAGCCATTTCTCTGGAGGAGCAGCAAAATTTCTCTGTCCTCAGCATTCACAGCC 540

Db 2580 CCCAGCCATTTCTCTGGAGGAGCAGCAAAATTTCTCTGTCCTCAGCATTCACAGCC 2639

QY 541 TCTCCCTGTGGATGACAGTGAGAAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600

Db 2640 TCTCCCTGTGGATGACAGTGAGAAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699

QY 601 CCCTGAGCTCAGCGGTACATCTCTGGAGCAGCCAGCGCGAGGCTCGAAGCTCCAGCTGGA 660

Db 2700 CCCTGAGCTCAGCGGTACATCTCTGGAGCAGCCAGCGCGAGGCTCGAAGCTCCAGCTGGA 2759

QY 661 ACATGCTGTGGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720

Db 2760 ACATGCTGTGGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2819

QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGCTCAAG 780

Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGCTCAAG 2879

QY 781 TGGGAGACATCGGCACCTGGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGTGCA 840

Db 2880 TGGGAGACATCGGCACCTGGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGTGCA 2939

QY 841 CCAAGAGCGGCGCCCTGTTCCGCTACGATGGAGTCCAGACTCCGGGCTACGACCTGC 900

Db 2940 CCAAGAGCGGCGCCCTGTTCCGCTACGATGGAGTCCAGACTCCGGGCTACGACCTGC 2999

QY 901 AGTGACACTGGGCCCTGTATGGGAGCTTTCGCTGGAGTGGAGGGTCAAGCATGGCCAGC 960





XX The invention relates to a method of classifying a breast cancer patient  
CC according to prognosis by determining the similarity between the level of  
CC expression of each of five genes for which markers are listed in the  
CC specification, in a cell sample taken from the breast cancer patient, to  
CC control levels of expression for each respective five genes to obtain a  
CC patient similarity value. The methods are useful for classifying a breast  
CC cancer patient according to prognosis. Kits and computer program products  
CC are useful for data analysis using the diagnostic, prognostic and  
CC statistical methods of the invention. This sequence corresponds to a  
CC marker used in the method of the invention.  
XX  
SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;  
  
Query Match 70.3%; Score 1850.2; DB 13; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;  
  
QY 1 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60  
DB 2100 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159  
  
QY 61 CTGCAGCGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
DB 2160 CTGCAGCGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219  
  
QY 121 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACGCCAATCAAGCCAAATTACC 180  
DB 2220 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACGCCAATCAAGCCAAATTACC 2279  
  
QY 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGCCCGGCCAG 240  
DB 2280 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGCCCGGCCAG 2339  
  
QY 241 CTGAGAGACAACAGGAGAGCCCTTAAGCTCAGAGCTCTCTCCCAACAGAGCCCGCAG 300  
DB 2340 CTGAGAGACAACAGGAGAGCCCTTAAGCTCAGAGCTCTCTCCCAACAGAGCCCGCAG 2399  
  
QY 301 AGCCAAACAAAGTCTCCTCCCTTGTACTTTGAGCAAGAGGAGTCTGGGATGTGGAAACCT 360  
DB 2400 AGCCAAACAAAGTCTCCTCCCTTGTACTTTGAGCAAGAGGAGTCTGGGATGTGGAAACCT 2459  
  
QY 361 TACCTCTGTCTCCTCTGGAGCAGCCCTTGCCAGAAACCCCAAGCTCACAGAGCGGAAAG 420  
DB 2460 TACCTCTGTCTCCTCTGGAGCAGCCCTTGCCAGAAACCCCAAGCTCACAGAGCGGAAAG 2519  
  
QY 421 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 480  
DB 2520 CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 2579  
  
QY 481 CCCAGCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 540  
DB 2580 CCCAGCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 2639  
  
QY 541 TCTCCTCTGTCGATGACAGTGAGAAACCCATCAAGGCCCTCTCAAAGCTCGCGGACA 600  
DB 2640 TCTCCTCTGTCGATGACAGTGAGAAACCCATCAAGGCCCTCTCAAAGCTCGCGGACA 2699  
  
QY 601 CCTTGAGCTCAGCGTACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 660  
DB 2700 CCTTGAGCTCAGCGTACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 2759  
  
QY 661 ACATGTGTGTGGCCGGGGCGGCCCAACGACACCCCAAGCTATTTCATGTGTGAAG 720  
DB 2760 ACATGTGTGTGGCCGGGGCGGCCCAACGACACCCCAAGCTATTTCATGTGTGAAG 2819  
  
QY 721 TCCAAATACAGTCTCTTAATGTGAACAACCTGCACATCCGGGAGTTCCAACGGGTCAAAG 780  
DB 2820 TCCAAATACAGTCTCTTAATGTGAACAACCTGCACATCCGGGAGTTCCAACGGGTCAAAG 2879  
  
QY 781 TGGGAGACATCGCACTCGCATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 840  
DB 2880 TGGGAGACATCGCACTCGCATCAGCAGCAGATCCAGCTGCAGCCTTCAGCTTGGTCA 2939

QY 841 CCAAAGACGGGCGACCTGTTTCGCTACGACATGAGAGTGCCAGACTCGGGCATCGACCTGC 900  
DB 2940 CCAAAGACGGGCGACCTGTTTCGCTACGACATGAGAGTGCCAGACTCGGGCATCGACCTGC 2999  
  
QY 901 AGTGCACACTGGGCCCCGTGATGAGCGAGCTTCGCTGGAGCTGAGAGGTCAAGCATGGCCAGC 960  
DB 3000 AGTGCACACTGGGCCCCGTGATGAGCGAGCTTCGCTGGAGCTGAGAGGTCAAGCATGGCCAGC 3059  
  
QY 961 TGGAGAACAGCCCTTAACCTCGCCCTCCACCGCGGCTCCACACTGCGCGAAACAGCGCT 1020  
DB 3060 TGGAGAACAGCCCTTAACCTCGCCCTCCACCGCGGCTCCACACTGCGCGAAACAGCGCT 3119  
  
QY 1021 TCCTGCTCGGTGACAGATGCTGCCCTGAAACACAGAGCTCAGCGCTTCCCAGGGGATVTG 1080  
DB 3120 TCCTGCTCGGTGACAGATGCTGCCCTGAAACACAGAGCTCAGCGCTTCCCAGGGGATVT - 3177  
  
QY 1081 NCCAGCCCCCGGCTCARGAGTGGGAACAGGCGCTCGNCAGCAGNAGNAGGTNGGGG 1140  
DB 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGCGCTCG-----CAGCGCAAGGTGGGG 3230  
  
QY 1141 CAAGCNAGAAATGCTCCAGGATTTTCAACNCTCAGCCCNCTGCCANCCCTGCTGAADA 1200  
DB 3231 GCAAGCAAAATGCTCCAGGATTTTCACTGAGCCCTGCCCCAC-----CTGCTGA 3284  
  
QY 1201 AAACAYTNCCGCCACGTGAAGAGACAGAGGAGGATGNCAGAGGTNNACCTTGGGGAA 1260  
DB 3285 AAAAACATCCGCCAGTGAAGAGACAGAGGAGATGSC---AGGAGTTACTTGGGGAA 3340  
  
QY 1261 ACAAAAAGGAGATCTTNTTCTGCGCCCTGCTCAGTNCAGTGGCTGACCTGACCCGCTGG 1320  
DB 3341 ACAAAAAGGAGATCTT--TTCTGCGCCCTGCTCCAGT--CGAGTTGGCCCTGA--CCGCTTG 3396  
  
QY 1321 ANTCAGTGACCATTTGTTGGCAGANCAAGGAGAGCAGCTTCCAGCCCTGGGTGAGAAGGG 1380  
DB 3397 GATCAGTGACCATTTGTTGGCAGA--CAGGAGAGCAGCTTCCAGCTTGGGTGAGAAGGG 3455  
  
QY 1381 GTGGGCGAGCCCTTGGGCCCTCAACCTNCCAGGCTGCTGTGNAGAGTCAAGTGTGTA 1440  
DB 3456 GTGGGCGAGCCCTTGGGCCCTCAACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGTA 3513  
  
QY 1441 AGGNCCCAAANCTCAGNTTTCAGTGACAGACAGGTCAGCAGGTATGCCGCCCGNTA 1500  
DB 3514 AGGGCCCAAACTC---AGGTTTCAGTGCAGAACACAGGT--CAGAGGTATGCCGCCCGTAG 3569  
  
QY 1501 GGTAAANNNGGGGGCCCTCTNAAAACCCCTTGCTNGGCCCTNCACTNGGCCAGCTCANCC 1560  
DB 3570 GTTAA----GGGGGCCCTCTAAAACCCCTTGCTGSCCTCA--CCTGSCCAGTCA--CCC 3621  
  
QY 1561 CTTTGGGTGTAGGGGAAAAAATGCTGACCTTGGGAAGGTGTCCTGGGTAGAAATACAC 1620  
DB 3622 CTTTGGGTGTAGGGGAAAAAATGCTGACCTTGGGAAGGT--CCCTGGGTAGAAATACAC 3680  
  
QY 1621 CACACTTTTCAGTGTGTTGCAACACAGGTCCGTGAGTTGACCTCTGCTTTCAGCAGAGGACC 1680  
DB 3681 CACACTTTTCAGTGTGTTGCAACACAGGTCCGTGAGTTGACCTCTGCTTTCAGCAGAGGACC 3740  
  
QY 1681 AAAGAAGGTGTGAAGTGAAGTGTTCAGTNCCTCCAGACATGTGCCCTTTGCTGCTG 1740  
DB 3741 AAAGAAGGTGTGAAGTGAAGTGTTCAGT--CCCCAGACATGTGCCCTTTGCTGCTG 3799  
  
QY 1741 GCTACCACTTCTCCAGAGCAGAGGCCCGAGCCCTTCAGGCCACAGCACTGCCCCAG 1800  
DB 3800 GCTACCACTTCTCCAGAGCAGAGGCCCGAGCCCTTCAGGCCACAGCACTGCCCCAG 3859  
  
QY 1801 ACTCGCTCGCACTCAGTTTCCCTCATCTCTGAAGGTGAAGGGTGTATGAGGATATGCTGA 1860  
DB 3860 ACTCGCTCGCACTCAGTTTCCCTCATCTCTGAAGGTGAAGGGTGTATGAGGATATGCTGA 3919  
  
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAGGNAAAGCAGCAGAGAGAGCY 1920  
DB 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGAGCG- 3976





```
QY 721 TCCAAATACAGTCTCTTAATGTGTGAACACTGTCACATCCGGAGTTCCACCGGGTCAAG 780
D 2870 TCCAAATACAGTCTCTTAATGTGTGAACACTGTCACATCCGGAGTTCCACCGGGTCAAG 2929
QY 781 TGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCA 840
D 2930 TGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTCGAGCTTTCAGCTTGGTCA 2989
QY 841 CAAAGACGGGAGCGCTGTCCTAGCAGATGGAGTGCAGATCGGGCATCGACCTGC 900
D 2990 CAAAGACGGGAGCGCTGTCCTAGCAGATGGAGTGCAGATCGGGCATCGACCTGC 3049
QY 901 AGTGACACTGGCCCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGTTCAAGCATGSCCAGC 960
D 3050 AGTGACACTGGCCCCCTGATGGCAGCTTCGCCCTGGAGCTGGAGGTTCAAGCATGSCCAGC 3109
QY 961 TGGAGAACAGGCCCTAACCCCTGCCCTCCACCCCGGCTCCACACTGCCGGAAGCAGCCT 1020
D 3110 TGGAGAACAGGCCCTAACCCCTGCCCTCCACCCCGGCTCCACACTGCCGGAAGCAGCCT 3168
QY 1021 TCTGCTCGGTGACAGTCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGATVYG 1080
D 3169 TCTGCTCGGTGACAGTCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGATVYG 1080
QY 1081 NCCAGCCCCCGGCTCAGCAGNTGGGAACAGGGCTTCGNCAGCAGCAGGAGTNGGGG 1140
D 3225 TCCAGCCCCCGGCTCAGCAGNTGGGAACAGGGCTTCGNCAGCAGCAGGAGTNGGGG 3277
QY 1141 CAAGCNAGATCCCTCCAGGATTTTCAACCTGAGCCCTGCCAGCCCTGCTGAADA 1200
D 3278 CAAGCNAGATCCCTCCAGGATTTTCAACCTGAGCCCTGCCAGCCCTGCTGAADA 1200
QY 1201 AAACAYNCCGACGCTGAAGAGACAGAGGAGTGNACAGGAGTTNACCTYGGGAA 1260
D 3335 AACACT--CCGCCACGTGNAGAGACA---GAGGAGATGGCAGGAGTTACCTCGGGAAA 3388
QY 1261 ACAAAACAGGATTTTTCGCCCTGCTTCAGTNCAGTGTGCTGACCCCTGG 1320
D 3389 CAACAGGATTTTTCGCCCTGCTTCAGTNCAGTGTGCTGACCCCTGG 3440
QY 1321 ANTCACTGACATTTTGTGGCAGANCAGGGAGACAGCTTCACGCTGGGTGAGAAGG 1380
D 3441 ATCACTGACATTTTGTGGCAGANCAGGGAGACAGCTTCACGCTGGGTGAGAAGG 3498
QY 1381 GTGGCGAGCCCTTCGGCCCTCACCCTNCCAGGCTGCTGTGNAGAGTCAAGTGTGTA 1440
D 3499 GTGGCGAGCCCTTCGGCCCTCACCCTNCCAGGCTGCTGTGNAGAGTCAAGTGTGTA 3556
QY 1441 AGGNCACAAANCTCAGNTTCAGTGCAGAACCCAGGTCAGAGGTATGCCCGCCGNTA 1500
D 3557 AGGNCACAAANCTCAGNTTCAGTGCAGAACCCAGGTCAGAGGTATGCCCGCCGNTA 3612
QY 1501 GGTAAANNNGGGGCTCTNAAACCCCTTGCTNGGCCCTNCCAGTTCAGCTCAGCC 1560
D 3613 GTTAA---GGGGGCTCTNAAACCCCTTGCTNGGCCCTNCCAGTTCAGCTCAGCC 3664
QY 1561 CTTTGTGGGTAGGGGAAAGAAATCCCTGACCCCTGGGAGGCTWCCTGGTGTAGATACAC 1620
D 3665 CTTTGTGGGTAGGGGAAAGAAATCCCTGACCCCTGGGAGGCTWCCTGGTGTAGATACAC 3723
QY 1621 CACACTTTTTCAGTGTGTCACACAGCTCTGAGTTGACCTCTGTTTCAGCCCAAGGACC 1680
D 3724 CACACTTTTTCAGTGTGTCACACAGCTCTGAGTTGACCTCTGTTTCAGCCCAAGGACC 3783
QY 1681 AAAGAGGTGTAGTCAAGTGGTTCATGNTCCAGACATGTCCTTCTGCTG 1740
D 3784 AAAGAGGTGTAGTCAAGTGGTTCATGNTCCAGACATGTCCTTCTGCTG 3842
QY 1741 GCTACCACTTCTCCACAGCAGCAGGCCCCGAGCCCTTCAGGCCAGCAGCTGCCCCAG 1800
D 3843 GCTACCACTTCTCCACAGCAGCAGGCCCCGAGCCCTTCAGGCCAGCAGCTGCCCCAG 3902
```

```
QY 1801 ACTGCTG3CACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGACAGGATATGCCCTGA 1860
D 3903 ACTGCTG3CACTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGACAGGATATGCCCTGA 3962
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTCTCTNAAAGGAAAGCAGCAGAGAGAGAG 1920
D 3963 CAGGAACAGTCTGTGGATGGACATGATCAGTCTCTNAAAGGAAAGCAGCAGAGAGAGAG 4020
QY 1921 TCCGGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCTGTGCTGTTNCCCGCAGNAGACA 1980
D 4021 TCCGGCGCCCGCAGNCCCACTNATCAGTGTGTTNCCCGCAGNAGACA 4069
QY 1981 GCTNCAGNATCANCACACTGACACTNACCTTCGCTGCTGCTGCTGCTGCTGCTGCTG 2040
D 4070 GCTNCAGNATCANCACACTGACACTNACCTTCGCTGCTGCTGCTGCTGCTGCTGCTG 4127
QY 2041 CCGNACGGCACTTTGCACTNCTGATGACCTTCAAGACACTTTTCATGCTGCTGCTGCTGCT 2100
D 4128 CCG-ACGGCACTTTGCACTNCTGATG-TCGTATG-ACCTCAAGACACTTTTCATGCTGCTGCTG 4184
QY 2101 GCAGGNCAGGNCAGGNCAGTGCACACTGACACTGACACTGACACTGACACTGACACTG 2160
D 4185 GCG-ACGGCACTTTGCACTNCTGATG-TCGTATG-ACCTCAAGACACTTTTCATGCTGCTGCTG 4234
QY 2161 GGTGNAAGGGGNCACAGTCTTGGCTGTCANCATGATGCTGCTGCTGCTGCTGCTGCTG 2220
D 4235 GGTGNA--GGGACACAGTCTTGGCTGTCANCATGATGCTGCTGCTGCTGCTGCTGCTG 4289
QY 2221 NCCAGNATTTCTTAAGATAGCANCCTTNCCTTNCCTTNCCTTNCCTTNCCTTNCCTT 2280
D 4290 --CCAGATTTCTTAAGATAGCANCCTTNCCTTNCCTTNCCTTNCCTTNCCTTNCCTT 4345
QY 2281 CCCAGGGGAGTACTCAGGACTCAGCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 2340
D 4346 CCCAGGGGAGTACTCAGGACTCAGCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 4403
QY 2341 TGTCTGTAGCTCAACCTCTCTGGGCGAGGGGACGCGAGACTCCGCTGGGAGAGAGCTCAT 2400
D 4404 TGTCTGTAGCTCAACCTCTCTGGGCGAGGGGACGCGAGACTCCGCTGGGAGAGAGCTCAT 4463
QY 2401 TCCACATCTTTGCAAGACAGCTTTTNGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
D 4464 TCCACATCTTTGCAAGACAGCTTTTNGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 4522
QY 2461 GGGGAGAGAGCCCGGCGCCCGCAGCACAATAAGACTGCGAGCTTGTGCTGCTGCTGCTG 2520
D 4523 GGGGAGAGAGCCCGGCGCCCGCAGCACAATAAGACTGCGAGCTTGTGCTGCTGCTGCTG 4582
QY 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTATGACAAATGTT 2572
D 4583 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTATGACAAATGTT 4634
```

## RESULT 8

ADQ67410

ID ADQ67410 standard; cDNA; 3152 BP.

XX AC ADQ67410;

XX AC ADQ67410;

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

XX DT 07-OCT-2004 (first entry)

ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;  
cytotoxic; gene therapy; diagnostic marker; morbid state; osteoporosis;  
neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
cancer.

Homo sapiens.

OS

XX

EP1440981-A2.

XX

28-JUL-2004.

FD

XX PF 21-JAN-2004; 2004EP-00001196.  
XX PR 21-JAN-2003; 2003JP-00102206.  
XX PR 09-MAY-2003; 2003JP-00131392.  
XX XX  
XX PA (REAS-) RES ASSOC.BIOTECHNOLOGY.  
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
XX PI Yamamoto J, Isono Y, Nagai K, Irie R;  
XX XX  
XX DR WPI: 2004-535376/52.  
XX DR P-PSDB; ADQ67717.  
XX XX  
XX PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,  
XX PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX XX  
XX PS Claim 1; SEQ ID NO 4571; 2449pp; English.  
XX XX  
XX CC The invention relates to 2495 novel polynucleotides (I) and their encoded  
XX CC polypeptides, sequences hybridizing to these nucleotides, sequences  
XX CC encoding partial polypeptides and sequences having 70% or 90% identity to  
XX CC the nucleotide and protein sequences. The nucleotides and polypeptides  
XX CC are useful as diagnostic markers or therapeutic target for the diseases  
XX CC or morbid states. They are also useful for treating osteoporosis,  
XX CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,  
XX CC dementia and various cancers. This sequence corresponds to a nucleotide  
XX CC sequence of the invention.  
XX XX  
XX SQ Sequence 3152 BP; 704 A; 1000 C; 874 G; 574 T; 0 U; 0 Other;

Query Match 67.3%; Score 1771.6; DB 12; Length 3152;  
Best Local Similarity 89.2%; Pred. No. 0;  
Matches 2287; Conservative 5; Mismatches 186; Indels 87; Gaps 35;

QY 1 CCCCTCTCACAGCCCGCCATCAAGAGGGCTTGAGGAAGAGCCCATCACCGGTGT 60  
DB 675 CCCCTCTCACAGCCCGCCATCAAGAGGGGTGAGGAAGAGCCCATCACCGGTGT 734  
QY 61 CTGACGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120  
DB 735 CTGACGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 794  
QY 121 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180  
DB 795 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 854  
QY 181 ACCAGCCCTTCATGCCAGCCGAGAGAGCTTTCCGCAAGGCCCCAGGGCCCGCCAG 240  
DB 855 ACCAGCCCTTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCGCCAGGGCCCGCCAG 914  
QY 241 CTGAGGAGACAACAGGCAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCGCCAG 300  
DB 915 CTGAGGAGACAACAGGCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCGCCAG 974  
QY 301 AGCCAAACAAAGTCTCTCTTGAATTTGACCAAGGAGGAGTCTGGGATGTGGGAACCT 360  
DB 975 AGCCAAACAAAGTCTCTCTTGAATTTGACCAAGGAGGAGTCTGGGATGTGGGAACCT 1034  
QY 361 TACCTCTGTCTCTCTGGAGCCAGCCCTCGCAGAAACCCAGCTACACGAGCGGAAAG 420  
DB 1035 TACCTCTGTCTCTCTGGAGCCAGCCCTCGCAGAAACCCAGCTACACGAGCGGAAAG 1094  
QY 421 CAAACGTCGGGAGAGGAGTGCAGAGCTGGGAATAGAAATATTCTCAACAGCCGTGT 480  
DB 1095 CAAACGTCGGGAGAGGAGTGCAGAGCTGGGAATAGAAATATTCTCAACAGCCGTGT 1154  
QY 481 CCCAGCCATTTCTCTGGAGGAGGAGGAGAAATTTCTCTGTGCTCAGCATCGACAGCC 540  
DB 1155 CCCAGCCATTTCTCTGGAGGAGGAGGAGGAGAAATTTCTCTGTGCTCAGCATCGACAGCC 1214  
QY 541 TCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAAGCTTCGCGGACA 600  
DB 1215 TCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAAGCTTCGCGGACA 1274  
QY 601 CCCTGAGCTCAGCGGTACACTCTCTGAGCAGAGCCAGGCGGAGGCTCGAAGCTCCAGCTCGA 660  
DB 1275 CCCTGAGCTCAGCGGTACACTCTCTGAGCAGAGCCAGGCGGAGGCTCGAAGCTCCAGCTCGA 1334  
QY 661 ACATGTGTGTCGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGTGAAG 720  
DB 1335 ACATGTGTGTCGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGTGAAG 1394  
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCCGGAGTTCCACCGGTCAAAG 780  
DB 1395 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCCGGAGTTCCACCGGTCAAAG 1454  
QY 781 TGGGAGACATCGCCACTTGGCATCAGCAGCCAGATCCAGCTTCAGCTTGGTCA 840  
DB 1455 TGGGAGACATCGCCACTTGGCATCAGCAGCCAGATCCAGCTTCAGCTTGGTCA 1514  
QY 841 CCAAAGACGGGAGCTGTTTGGCTACGATGAGAGTGCAGACTCGGGCATCGACCTGC 900  
DB 1515 CCAAAGACGGGAGCTGTTTGGCTACGATGAGAGTGCAGACTCGGGCATCGACCTGC 1574  
QY 901 AGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGAGGCTCAAGCATGGCCAGC 960  
DB 1575 AGTGCACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGAGGCTCAAGCATGGCCAGC 1634  
QY 961 TGCAGAACAGGCCCTAACCCCTGCCCTCCACGGCGGCTCCACACTGCGG-AAGCAGCCT 1020  
DB 1635 TGCAGAACAGGCCCTAACCCCTGCCCTCCACGGCGGCTCCACACTGCGG-AAGCAGCCT 1693  
QY 1021 TCTGTCTGGTGCACGATGCTGCCCTGAAAAACAGACTCAGCCGTTCCAGGGGATYTG 1080  
DB 1694 TCTGTCTGGTGCACGATGCTGCCCTG-AACACAGGCTCAGCCGTTCCAGGG-ATC 1749  
QY 1081 NCAGCCCCCGGCTCARCAGNTGGGACACAGGGCTTCGNCAGCAGCAGNAGTNGGGG 1140  
DB 1750 TGCAGCCCCCGGCTCAGCAGTGGGACACAGGGCTTCGACAGCAGCAAG-----GTGGGG 1802  
QY 1141 CAAGCAGAAATGCTCTCCAGGATTTTACANCTGAGCCNTGCCCANCTCTGCTGAADA 1200  
DB 1803 GCAAGCAGATGCTCTCCAGGATTTTACACC---TGAGCCCTGCCACCTCTGTGAGAA 1859  
QY 1201 AAACATYNNCCGCACTGTAAGAGACAGAGAGGATGAGNAGGAGTTNNACTYVGGGAA 1260.  
DB 1860 AACACT--CGCCACCTGTAAGAGACA---GAGGAGGATGGCAGGAGTTACCTCGGAAA 1913  
QY 1261 ACAAACAGGAGTCTTNTTCTGCCCCCTGCTCCAGTNCAGTTGGCCCTGNAACCCGCTTG 1320.  
DB 1914 CAAACAGGATCTTCTCTGCCCCCTGCTCCAG-----TCGAGTTGGCCCTGACCCGCTT--G 1964  
QY 1321 ANTCAAGTACCATTTGTTGGCAGANCAAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGGG 1380  
DB 1965 GATCAGTGACATTTGCTGGCAGA-CAGGGGAGAGCAGCTTCCAGCTGGGTCAAGAGGG 2023  
QY 1381 GTGGGAGAGCCCTTCGGCCCCCTCACCCCTNCCAGGCTGCTGTGAGAGTGTCAAGTGTGTA 1440  
DB 2024 GTGGGAGAGCCCTTCGGCCCCCTCACCCCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGTA 2081  
QY 1441 AGGNCCTCAANCTCAGGNTTCAAGTGCAGAACCAAGTNCAGCAGGATATGCCCCGCCGNTA 1500  
DB 2082 AGGGCCCCAACTC---AGGTTCAAGTGCAGAACCAAGGT-CAGCAGGATATGCCCCGCCGTAG 2137  
QY 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTCCCTNGGCTNCACTTCGCGCAGCTCANCCC 1560  
DB 2138 GTTAA---GGGGGCCCTCTAAACCCCTTGCCTGGCCCTCA---CCTGGCAGCTCA-CCC 2189  
QY 1561 CTTTGTGGGTAGGGGAAAAGAAATGCTGACCCCTGGGAAGGCTTCCCTGGTGTAGTAATACAC 1620  
DB 2190 CTTTGTGGGTAGGGGAAAAGAAATGCTGACCCCTGGGAAGGCT-CCCTGGTGTAGTAATACAC 2248  
QY 1621 CACATTTTTCAGGTTGTGTGCAACACAGGTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 1680  
DB 2249 CACATTTTTCAGGTTGTGTGCAACACAGGTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 2308

QY	1681	AAAGAGGTGTGAAGTCAAGTGGTTCTCAGTNC	CCCGAGACATGTGCCCTTTGCTGCTG	1741
DB	2309	AAAGAGGTGTGAAGTGAAGTGGTTCTCAGTNC	CCCGAGACATGTGCCCTTTGCTGCTG	2367
QY	1741	GCTACCACTCTTCCCGAGAGCAGAGCCCGAGC	CCCTTCAGGCCCGAGCACTGCCCCAG	1800
DB	2368	GCTACCACTCTTCCCGAGAGCAGAGCCCGAGC	CCCTTCAGGCCCGAGCACTGCCCCAG	2427
QY	1801	ACTCGTGGCACTCAGTTCCTCATCTGTGTAAAG	TGAAGGTGATGAGGATATGCTCTGA	1860
DB	2428	ACTCGTGGCACTCAGTTCCTCATCTGTGTAAAG	TGAAGGTGATGAGGATATGCTCTGA	2487
QY	1861	CAGGAACAGTCTGTGGATGGACATGATCAGTCT	TAAGNAAAGCAGAGAGAGAGCGY	1920
DB	2488	CAGGAACAGTCTGTGGATGGACATGATGATGCTAA	AGCAGAGAGAGAGAGCGC	2545
QY	1921	TCCGGCGCCCGAGNCCCGCACTNATCAGTGTN	CCAGCGTGTCTGCTGCCAGNAGCACA	1980
DB	2546	TCCGGCGCCCGAGNCCCGCACTNATCAGTGTN	CCAGCGTGTCTGCTGCCAGNAGCACA	2594
QY	1981	GCTNCAGNATCANCACTGACATNCACCTTNGCC	TGCCCTTNGGCCAANGAGGTACTG	2040
DB	2595	GCACAGTTCAGCATCAGTGCACACTCACCTTGC	CTGCCCTTNGGCCAANGAGGTACTG	2652
QY	2041	CCGNAGCGCACTTTGCACTCTGATGACACTTCA	AGCAGTTCATGCTGCTGCTCTTNG	2100
DB	2653	CCG-ACGCACTTTGCACTCTGATGACACTTCA	AGCAGTTCATGCTGCTGCTCTTNG	2709
QY	2101	GCAGGNCAGGNCAGGNCAGTGCACANCTGTAG	GNAGCATANGCAANGCCAGGATGG	2160
DB	2710	GGCGAGGCGAGGCGAGTGCACACTGT-----	AGGAGCATAGCAAGCCAGGATGG	2759
QY	2161	GGTGNAAGGGANCAAGTCTTTGAGCTGTCCAN	CATGATGTGACTNCTCAAACTCTTN	2220
DB	2760	GGTGAAGGG--ACACAGTCTTGAGCTGTCCA-	CATGATGTGACT-CCTCAAACTCTT-	2814
QY	2221	NCCAGNATTTCTTAAGATAGCANCCCCCTTNC	CCCCCATTTGCCCGACTTAGCTCTTCT	2280
DB	2815	--CCAGATTTCTTAAGATAGCACCCCC--TT	CCCCCATTTGCCCGACTTAGCTCTTCT	2870
QY	2281	CCCAGGGAGCTANCTCAGGACTCACGTAGCAT	TAAATCAGCTGTGNAATCGTCAGGGG	2340
DB	2871	CCCAGGGAGGCTA-CTCAGGACTCACGTAGCAT	TAAATCAGCTGTG-AATCGTCAGGGG	2928
QY	2341	TGCTCTGCTAGCTCAACTCTCTGGGCGAGGGG	ACGCCAGACTCCGTGGGAGAGCTCAT	2400
DB	2929	TGCTCTGCTAGCTCAACTCTCTGGGCGAGGGG	ACGCCAGACTCCGTGGGAGAGCTCAT	2988
QY	2401	TCCACATCTTGCCAAAGACAGCTTTTNGTCC	AGCTGTCCACATTTGAGTCAAGTCTCCC	2460
DB	2989	TCCACATCTTGCCAAAGACAGCTTTTNGTCC	AGCTGTCCACATTTGAGTCAAGTCTCCC	3047
QY	2461	GGGAGAGAGCCCCGGCCCCCAGCACATAAGA	AACTGCAGCTTGGTACTGCAGAGTCTG	2520
DB	3048	AGGAGAGAGCCCCGGCCCCCAGCACATAAGA	AACTGCAGCTTGGTACTGCAGAGTCTG	3107
QY	2521	GGTTGTAGAGACTCTTTGTGAAGCAATAAAG	TTTGGGTGATGAC	2565
DB	3108	GGTTGTAGAGACTCTTTGTGAAGCAATAAAG	TTTGGGTGATGAC	3152

## RESULT 9

AAV73917

AAV73917  
ID AAV73917 standard; CDNA; 3156 BP.

XX  
XX

AC AAV73917;

XX

DT 04-MAR-1999 (first entry)

XX

DE Human NIK cDNA.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

KW NIK; Nuclear factor-kappa B; NF-kB; signal transduction; TNF; human;

tumour necrosis factor; NF- $\kappa$ B-inducing kinase; screening; interaction; modulator; immune response; inflammatory response; viral gene; diagnostic; therapy; ds.

**Homo sapiens.**

US5854003-A.

29-DEC-1998.

26-FEB-1998. 98TIS-00032475

03-.III.-1997. 97IIS-00887518

(TITT.A- ) TITT.APTK TNC

W. Y. Botch M.

WDT: 1000 004003/00

Screening agents for modulating interaction of nuclear factor kappaB inducing kinase - with kinase-binding target, useful for controlling levels of the kinase, for treatment and diagnosis of conditions associated with e.g. inhibition of signal transduction by tumour necrosis factor.

Disclosure: Col 11-14: 16pp: English

This sequence encodes a novel human nuclear factor-kappaB (NFkB)-inducing kinase (NIK) which is used in a method for screening for agents that modulate the interaction of NIK with a NIK-binding target. The encoded protein can be used as a modulator of cellular functions at the NIK level, or for development of such compounds. NFkB is involved in expression of many immune and inflammatory responses and of some important viral genes. The protein may be used diagnostically and therapeutically, in conditions associated with abnormal utilisation of pathways that involve NFkB, e.g. inhibition of signal transduction by tumour necrosis factor (TNF).

Sequence 3156 BP: 736 A: 991 C: 895 G: 534 T: 0 U: 0 Other:

43.18: Score 1133.4: DB 2: Length 3156:

1st Local Similarity 94.8%: Pred. No. 2: 1e-305: SCOE 1133:4; DB 43.1%;

4: Mismatches 43; Indels 21; Gaps 7;  
 4: Mismatches 43; Indels 21; Gaps 7;

1 CCCCTCTCAGCCCAGGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTG 60

1868 CCCCTCTCAGGCCAGGCCATCCAAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 1927

61 CTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCCTGAAGA 120

1928 CTGCA GCGGAGCTGCGGACCGGACGCTGACACCGGGCCTACACGCAACGTGGGACGCTCTGACACG 1987

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

\_\_\_\_\_

\_\_\_\_\_

[illegible]

# THE UNIVERSITY OF CHICAGO

[illegible]

```
Qy 421 CAACGCTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATATTCCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACCTGACAGAGCTGGAAATAGAAATATTCCTCAACAGCCTGT 2347
Qy 481 CCCAGCCATTTCTCTGAGGAGCAGGAGCAAAATTCCTGCTGCTCAGAGCTCAGACAGCC 540
Db 2348 CCCAGCCATTTCTCTGAGGAGCAGGAGCAAAATTCCTGCTGCTCAGAGCTCAGACAGCC 2407
Qy 541 TCTCCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2467
Qy 601 CCCTGAGCTCAGGCTTACATCTCTGAGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTTGA 660
Db 2468 CCCTGAGCTCAGGCTTACATCTCTGAGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTTGA 2527
Qy 661 ACATGGTGTGCGCGCGCGCGCCACCCAGCACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2528 ACATGGTGTGCGCGCGCGCGCCACCCAGCACCCCAAGCTATTTCAATGGTGTGAAG 2587
Qy 721 TCCAATACAGTCTCTTAATGTTGAACACCTGACATCCCGGAGTTCCACCGGCTCAAG 780
Db 2588 TCCAATACAGTCTCTTAATGTTGAACACCTGACATCCCGGAGTTCCACCGGCTCAAG 2647
Qy 781 TGGAGACATCGCACCTGGGATCAGCAGCCAGATCCAGCTGACGCTTCAAGTTGTGCA 840
Db 2648 TGGAGACATCGCACCTGGGATCAGCAGCCAGATCCAGCTGACGCTTCAAGTTGTGCA 2707
Qy 841 CCAAGACGGCAGCTGTTGCTTACGACATGAGGTTCCAGACTCGGGCATCGACCTGC 900
Db 2708 CCAAGACGGCAGCTGTTGCTTACGACATGAGGTTCCAGACTCGGGCATCGACCTGC 2767
Qy 901 AGTGACACTGGCCCTGATGGAGCTTGGCTGAGCTGAGGCTCAAGCATGGCCAGC 960
Db 2768 AGTGACACTGGCCCTGATGGAGCTTGGCTGAGCTGAGGCTCAAGCATGGCCAGC 2827
Qy 961 TGGAGACAGCCCTAACCTCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020
Db 2828 TGGAGACAGCCCTAACCTCCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 2887
Qy 1021 TCCTGCTCGTGACGATGCTGCTTGAACACAGGCTCAGCGTTCCAGGAGGATYTG 1080
Db 2888 TCCTGCTCGTGACGATGCTGCTTGAACACAGGCTCAGCGTTCCAGGAGGATYTG 2945
Qy 1081 NCCAGCCCCCGCTCARGAGTGGGAACAGGCGCTCGNACAGNACNAAGTNGGGG 1140
Db 2946 GCCAGCCCCCGCTCARGAGTGGGAACAGGCGCTCGNACAGNACNAAGTNGGGG 2998
Qy 1141 CAAGCAGAAATGCTCCAGGATTTACAGCTGAGCCCTGCCCCAGC-----CCTGCTGA 1200
Db 2999 GCAAGCAGAAATGCTCCAGGATTTACAGCTGAGCCCTGCCCCAGC-----CCTGCTGA 3052
Qy 1201 AAACATYTCGACGCTGAGAGACAGAGAGATGAGGATGAGGTTNACCTYGGGAA 1260
Db 3053 AAACATYTCGACGCTGAGAGACAGAGAGATGAGGATGAGGTTNACCTYGGGAA 3108
Qy 1261 ACAAACAGGATCTTTTCTGCTGCTGCTTCCAGTNCAGTGGGCTG 1309
Db 3109 ACAAACAGGATCTTTTCTGCTGCTGCTTCCAGTNCAGTGGGCTG 3155
```

## RESULT 10

AAV71603

ID AAV71603 standard; cDNA; 3156 BP.

XX AC

XX AC

DT 04-FEB-1999 (first entry)

XX AC

DE Human NF-kB-inducing kinase (NIK) polypeptide encoding cDNA.

XX NIK; IkB; TNF; kinase; nuclear factor kappa B; inhibition;

KW tumour necrosis factor; binding; genetic hybridisation; screening;

KW tumour necrosis factor; binding; genetic hybridisation; screening;

signal transduction; biopharmaceutical; human; ds.

Homo sapiens.

Key Location/Qualifiers  
CDS 1..2844/\*tag= a  
/product= "NIK polypeptide"

XX US5843721-A.

XX 01-DEC-1998.

XX 03-JUL-1997; 97US-00887518.

XX 03-JUL-1997; 97US-00887518.

XX (TULA-) TULARIK INC.

XX Wu L, Rothe M;

XX WPI; 1999-044580/04.

XX P-PSDB; AAW81564.

XX Probe, vector or recombinant nucleic acid encoding a polypeptide, especially human nuclear factor kappa-B-inducing kinase protein - useful for producing recombinant protein.

XX Claim 1; Col 11-14; 15pp; English.

XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK) polypeptide. The NIK polypeptide or its fragment has one or more activities selected from kinase activity and inhibitory activity; IkB kinase-alpha and beta binding activity and binding inhibitory activity; tumour necrosis factor (TNF) receptor-associated factor 2 binding activity and binding inhibitory activity; IkB binding activity and binding inhibitory activity; NF-kB activating and inhibitory activity. A vector containing the NIK nucleic acid can be used to transform host cells for the recombinant production of the protein. The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g. genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK kinase inhibitors to inhibit TNF signal transduction), and in the biopharmaceutical industry

XX Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 43.1%; Score 1133.4; DB 2; Length 3156;

Best Local Similarity 94.8%; Pred. No. 2.1e-305;

Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

Qy 1 CCCTCTCACAGCCCGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT 60

Db 1868 CCCTCTCACAGCCCGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT 1927

Qy 61 CTCAGCGAGCTGGAGGAAAGGTGAAACCGGCGCATACAGCAAGTGGAGGTCTGAAGA 120

Db 1928 CTCAGCGAGCTGGAGGAAAGGTGAAACCGGCGCATACAGCAAGTGGAGGTCTGAAGA 1987

Qy 121 GCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACGCCCAATCAAGCCATACC 180

Db 1988 GCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACGCCCAATCAAGCCATACC 2047

Qy 181 ACCAGACCTCTCCATGCCCGGAGAGAGCTTTTCGCCAAGGGGCCCAAGGGCCCGGCGAG 240

Db 2048 ACCAGACCTCTCCATGCCCGGAGAGAGCTTTTCGCCAAGGGGCCCAAGGGCCCGGCGAG 2107

Qy 241 CTGAGGAGCAACAGGCGAGAGCCCTTAAGTCCAGCTCTCTCCACAGAGCCCGGAG 300

Db 2108 CTGAGGAGCAACAGGCGAGAGCCCTTAAGTCCAGCTCTCTCTCCACAGAGCCCGGAG 2167

Qy 301 AGCCAAACAGTCTCTCTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 360

Db 2168 AGCCAAACAGTCTCTCTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 2227

QY	361	TACCTCTGCTCTCCCTGTGAGCGAGCCCTCTGCAGAAACCCAGTCTACCGAGCGGAAG	420
Db	2228	TACCTCTGCTCTCCCTGTGAGCGAGCCCTCTGCAGAAACCCAGTCTACCGAGCGGAAG	2287
QY	421	CAACCGTCCCGAGCAGAGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAACGCTGT	480
Db	2288	CAACCGTCCCGAGCAGAGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAACGCTGT	2347
QY	481	CCCAGCCATTTTCTCTGAGGAGCAGGACAAATTTCTCTGTGCTCAGCATCGACGCC	540
Db	2348	CCCAGCCATTTTCTCTGAGGAGCAGGACAAATTTCTCTGTGCTCAGCATCGACGCC	2407
QY	541	TCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAGGCTCTCAAAAGCTTCGGGGACA	600
Db	2408	TCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAGGCTCTCAAAAGCTTCGGGGACA	2467
QY	601	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGA	660
Db	2468	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGA	2527
QY	661	ACATGCTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG	720
Db	2528	ACATGCTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG	2587
QY	721	TCCAAATACAGTCTCTTAATGGTGAAACACTGTGCAATCCGGGAGTTCACCGGGTCAAAG	780
Db	2588	TCCAAATACAGTCTCTTAATGGTGAAACACTGTGCAATCCGGGAGTTCACCGGGTCAAAG	2647
QY	781	TGGGAGACATCGCCACCTGGCATCAGCAGCAGATCCCGAGCTGCAGCTTCAGCTTGTGCA	840
Db	2648	TGGGAGACATCGCCACCTGGCATCAGCAGCAGATCCCGAGCTGCAGCTTCAGCTTGTGCA	2707
QY	841	CCAAAGACGGGACGCTGTTGCTTACGACATGGAGGTGCCAGATCTCGGGCATCGACCTGC	900
Db	2708	CCAAAGACGGGACGCTGTTGCTTACGACATGGAGGTGCCAGATCTCGGGCATCGACCTGC	2767
QY	901	AGTGCACTATGGCCCTGTATGGCAGCTTTCGCTGAGAGGTGCCAGACTCGGCGCATCGACCTGC	960
Db	2768	AGTGCACTATGGCCCTGTATGGCAGCTTTCGCTGAGAGGTGCCAGACTCGGCGCATCGACCTGC	2827
QY	961	TGGAGAACAGGCGCTTAACCTTCGCCCTCCACCGCGGCTCCACATGCGCGGAAAGCAGCT	1020
Db	2828	TGGAGAACAGGCGCTTAACCTTCGCCCTCCACCGCGGCTCCACATGCGCGGAAAGCAGCT	2887
QY	1021	TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCGGTTCCCGAGGGGATYTG	1080
Db	2888	TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCGGTTCCCGAGGGGATYTG	2945
QY	1081	NCAGCCCCCGGCTCARGAGNTGGGAAACAGGGCTTCGNCAGNAGNAGNTGGGGG	1140
Db	2946	GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGGCTTCG--CAGCAAGGTTGGG	2998
QY	1141	CAAGCAGAAATGCTCCCGAGGATTTACANCTGAGCCNTGCCCCANCCCTCTGAADA	1200
Db	2999	GCAAGCAGAAATGCTCCCGAGGATTTACANCTGAGCCNTGCCCCANCCCTCTGAADA	3052
QY	1201	AAACAYTNCGCCACGTGAAGACAGAGGAGATGNCAGGAGTTNNACCTTGGGAA	1260
Db	3053	AAACAYTNCGCCACGTGAAGACAGAGGAGATGNCAGGAGTTNNACCTTGGGAA	3108
QY	1261	ACAAAACAGGATCTTTNTTCTGCCCTGTCTCCAGTNCAGTGTGGCGTG	1309
Db	3109	ACAAAACAGGATCTTTT-TTCTGCCCTGTCTCCAGT-CCAGT-TGGGCTG	3155

RESULT 11

RESUL I.  
AAV69285

AAV69263  
ID AAV69285 standard; cDNA; 3156 BP.

XX  
 09769AHY 01

AAV69285;

XX  
CONTENTS

DT 04-FEB-1999 (first entry)

Human NF- $\kappa$ B-inducing kinase (NIK) polypeptide encoding cDNA.

NIK; I $\kappa$ B; NF- $\kappa$ B; TNF; kinase; nuclear factor kappa B; inhibition; tumour necrosis factor; binding; genetic hybridisation; screening; signal transduction; biopharmaceutical; immunogen; pharmacological; transcription regulator; human; ds.

**Homo sapiens.**

Key	Location/Qualifiers
CDS	1 2844

```

/*tag= a
/product= "NIK polypeptide"

```

US5844073-A.

01-DEC-1998.

13-FEB-1998: 98US-00023321-

03-.III.-1997. 97118-00987518

(ATTN: A - ) THE ARMY THE

Mr. T. P. Matheson.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

P~PSDB; AAW81561.

New isolated peptide comprising a specified 947 amino acid sequence - has e.g. kinase activity, kinase inhibitory activity, IκB kinase-α binding activity, and IκB kinase-α binding inhibitory activity.

Disclosure; Col 11-14; 15pp; English.

This cDNA encodes a nuclear factor kappa B (NF- $\kappa$ B)-inducing kinase (NIK) polypeptide. The NIK polypeptide or its fragment has one or more activities selected from kinase activity and inhibitory activity; I $\kappa$ B kinase- $\alpha$  and beta binding activity and binding inhibitory activity; tumour necrosis factor (TNF) receptor-associated factor 2 binding activity and binding inhibitory activity; I $\kappa$ B binding activity and binding inhibitory activity, NF- $\kappa$ B activating and inhibitory activity. The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g. genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK kinase inhibitors to inhibit TNF signal transduction), and in the biopharmaceutical industry (e.g. as immunogens, reagents for isolating other transcription regulators, and reagents for screening chemical libraries for pharmacological agents).

Sequence 3156 BP: 736 A: 991 C: 895 G: 534 T: 0 U: 0 Other:

Query Match	43.1%;	Score 1133.4;	DB 2;	Length 3156;
Best Local Similarity	94.8%;	Pred. No. 2.1e-305;		
Other 1241	Concentrations	43		
	43	Matched		
	43	Matched		

Qy	1	CCCTCTCTCAGAGCCCGAGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT	60
Db	1868	CCCTCTCTCAGAGCCCGAGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT	1927
Qy	61	CTCAGCGGAGCTGGAGGGGAAGTGAACCGGSCACTACAGCAAGTGGGAGGTCTGAGA	120
Db	1928	CTCAGCGGAGCTGGGNGGGAAGTGNACCGGSCACTACAGCAAGTGGGAGGTCTGAGA	1987
Qy	121	GCCTTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCMAATCAAGCCMAATTACC	180
Db	1988	GCCTTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCMAATCAAGCCMAATTACC	2047
Qy	181	ACCAGACCTTCCATGCCACGCGAGAGAGCTTTGCCCCAAGGGCCCCCAGGGCCCGGGCCAG	240
Db	2048	ACCAGACCTTCCATGCCACGCGAGAGAGCTTTGCCCCAAGGGCCCCCAGGGCCCGGGCCAG	2107
Qy	241	CTGAGAGAGCAACAGGCAGAGCCCTTAAGTCTCAGAGCTCTCTCTCCACACAGAGCCCCCAG	300

Db 2108 CTGAGGACAAACAGGACAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 2167  
Qy 301 AGCCAAACAAAGTCTCTCCCTTGATCTTTGACAAAGGAGGAGTCTGGATGTGGAAACCT 360  
Db 2168 AGCCAAACAAAGTCTCTCCCTTGATCTTTGACAAAGGAGGAGTCTGGATGTGGAAACCT 2227  
Qy 361 TACCTCTGTCTCTCTCTGAGGACAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAG 420  
Db 2228 TACCTCTGTCTCTCTCTGAGGACAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAG 2287  
Qy 421 CAACCGTCCCGAGCAGCACTGACAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 480  
Db 2288 CAACCGTCCCGAGCAGCACTGACAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 2347  
Qy 481 CCCAGCCATTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCCAGAGCC 540  
Db 2348 CCCAGCCATTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCCAGAGCC 2407  
Qy 541 TCTCTCTGTGGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACA 600  
Db 2408 TCTCTCTGTGGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACA 2467  
Qy 601 CCCTGAGCTCAGGCGTACACTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTTCCAGCTGGA 660  
Db 2468 CCCTGAGCTCAGGCGTACACTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTTCCAGCTGGA 2527  
Qy 661 ACATGCTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTCAATGCTGTGAAAG 720  
Db 2528 ACATGCTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTCAATGCTGTGAAAG 2587  
Qy 721 TCCAAATACAGTCTTTAATGCTGAACACCTGACATCCGGGAGTTCACCGGGTCAAAG 780  
Db 2588 TCCAAATACAGTCTTTAATGCTGAACACCTGACATCCGGGAGTTCACCGGGTCAAAG 2647  
Qy 781 TGGGAGACATCGCCACTGCGATCAGCAGCCAGATCCCGAGCTCGAGCCTTCAGCTTGCTCA 840  
Db 2648 TGGGAGACATCGCCACTGCGATCAGCAGCCAGATCCCGAGCTCGAGCCTTCAGCTTGCTCA 2707  
Qy 841 CCAAAGCGGCGACCTGTTGCTTACACATGAGAGTGCAGAGTCCGGGATCGACCTGC 900  
Db 2708 CCAAAGCGGCGACCTGTTGCTTACACATGAGAGTGCAGAGTCCGGGATCGACCTGC 2767  
Qy 901 AGTGCACTGCGCCCTGATGGCAGCTTTCGCTGAGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 2768 AGTGCACTGCGCCCTGATGGCAGCTTTCGCTGAGAGCTGGAGGTCAAGCATGGCCAGC 2827  
Qy 961 TGGAGAACGCGCTTACCTGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 1020  
Db 2828 TGGAGAACGCGCTTACCTGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT 2887  
Qy 1021 TCTGCTCGGTGACAGATGCTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG 1080  
Db 2888 TCTGCTCGGTGACAGATGCTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATYT-- 2945  
Qy 1081 NCAGACCCCGGCTCAGCAGTGGGAAACAGGCGCTTCGACAGCAGCAGTTCAGGGGAT-- 1140  
Db 2946 GCCAGCCCGGCTCA--CAGTGGGAAACAGGCGCTCG----CAGCAGCAAGTGGGG 2998  
Qy 1141 CAACGAGATGCTCTCCAGATTTTACANCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 2999 GCAAGCAGATGCTCTCCAGATTTTACACTGAGCCTTCCAGCCTTCCAGCCTTCCAGCCT 3052  
Qy 1201 AAACATYTCGCCACGCTGAAGAGACAGAGGAGATGAGGATGAGGATGAGGATGAGGATGAGG 1260  
Db 3053 AAACATYTCGCCACGCTGAAGAGACAGAGGAGATGAGGATGAGGATGAGGATGAGGATGAGG 3108  
Qy 1261 AAAAAACAGGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1309  
Db 3109 AAAAAACAGGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3155

RESULT 12  
AAx87842

ID XX AAX87842 standard; DNA; 2844 BP.  
XX AC AAX87842;  
DT XX 09-NOV-1999 (first entry)  
DE XX NF-kB inducing kinase (NIK) DNA.  
XX NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;  
KW interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;  
KW anti-apoptotic; human; ss.  
OS Homo sapiens.  
XX WO9943704-A1.  
PN 02-SEP-1999.  
XX 25-FEB-1999; 99WO-US004110.  
XX 27-FEB-1998; 98US-0076299P.  
XX (REGC ) UNIV CALIFORNIA.  
PI Greene WC, Lin X, Gelezuinas R;  
XX WPI; 1999-518837/43.  
DR P-PSDB; AAY31665.  
XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent  
PT immune response, e.g. an inflammatory response or an anti-apoptotic  
PT response.  
XX Disclosure; Fig 10; 48pp; English.  
XX This DNA sequence codes for human wild-type NIK (see AAY31665), a NF-kB  
CC inducing kinase having serine/threonine kinase activity. The invention  
CC relates to: (a) an N-terminus deletion mutant NIK protein (see AAY31666);  
CC and (b) a kinase deficient NIK mutant protein (see AAY31667 and AAY31668)  
CC that inhibits auto-phosphorylation or transphosphorylation. The invention  
CC provides the molecular basis for cytokine induction of NF-kB-dependent  
CC immune and inflammatory responses, emphasising a role for both NIK-NIK  
CC and NIK-IKK (IKK-specific kinase) interactions. A novel and highly  
CC specific method for modulating NF-kB-dependent immune, inflammatory and  
CC anti-apoptotic responses is based on interruption of the critical protein  
CC interaction of NIK and IKK. The mutant NIK proteins are used in claimed  
CC methods for inhibiting NF-kB-dependent gene expression. The kinase-  
CC deficient NIK mutant proteins inhibit activation of IKK. The N-terminal  
CC deletion NIK mutant proteins bind to IKK, thus inhibiting NIK/IKK  
CC interaction  
XX Sequence 2844 BP; 665 A; 887 C; 809 G; 483 T; 0 U; 0 Other;

Query Match 37.1%; Score 977; DB 2; Length 2844;  
Best Local Similarity 100.0%; Pred. No. 9.9e-262;  
Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCCTCTCACAGCCAGGCGCATCCAGAGGGGCTGAGGAAGAGCCCATCCAGCGCTGT 60  
Db 1868 CCCTCTCACAGCCAGGCGCATCCAGAGGGGCTGAGGAAGAGCCCATCCAGCGCTGT 1927  
Qy 61 CTGACGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
Db 1928 CTGACGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 1987  
Qy 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCAATATACC 180  
Db 1988 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCAATATACC 2047  
Qy 181 ACCAGACCTTCCATGCCAGCGGAGAGCTTTCGCCAAGGGGCCCGAGGGCCCGGCCAG 240  
Db 2048 ACCAGACCTTCCATGCCAGCGGAGAGCTTTCGCCAAGGGGCCCGAGGGCCCGGCCAG 2107



QY 241 CTGAGAGACAACAGCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCCAG 300  
DB 2108 CTGAGAGACAACAGCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCCAG 2167  
QY 301 AGCCAAACAGTCTCTCTCTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTT 360  
DB 2168 AGCCAAACAGTCTCTCTCTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTT 2227  
QY 361 TACCTCTGTCTCTCTCTGAGCAGCCCTGAGCAAGCCCTGAGCAAGCCCTGAGCAAG 420  
DB 2228 TACCTCTGTCTCTCTCTGAGCAGCCCTGAGCAAGCCCTGAGCAAGCCCTGAGCAAG 2287  
QY 421 CAACCTCTCCGAGCAGGAACTGAGCAGCTGGAATAGAAATATTTCTCAACAGCTGT 480  
DB 2288 CAACCTCTCCGAGCAGGAACTGAGCAGCTGGAATAGAAATATTTCTCAACAGCTGT 2347  
QY 481 CCCAGCATTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTCTGAGCATCGACAGCC 540  
DB 2348 CCCAGCATTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTCTGAGCATCGACAGCC 2407  
QY 541 TCTCTCTGTCTGATGACGTGAGAGAACCCATCAAGGCTCTCAAGCTCTCGCGGACA 600  
DB 2408 TCTCTCTGTCTGATGACGTGAGAGAACCCATCAAGGCTCTCAAGCTCTCGCGGACA 2467  
QY 601 CCTGAGCTCAGGCTTACACTCTCTGAGCAGCAGCCAGGCTCGAAGCTCCAGCTGGA 660  
DB 2468 CCTGAGCTCAGGCTTACACTCTCTGAGCAGCAGCCAGGCTCGAAGCTCCAGCTGGA 2527  
QY 661 ACATGGTGTCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720  
DB 2528 ACATGGTGTCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2587  
QY 721 TCCAAATACAGTCTCTTAATGTGTGAACACTCTGCATCCGGAGTTCCACCGGGTCAAG 780  
DB 2588 TCCAAATACAGTCTCTTAATGTGTGAACACTCTGCATCCGGAGTTCCACCGGGTCAAG 2647  
QY 781 TGGGAGACATCGCCACTGSCATCAGCAGCAGATCCAGCTGCGACCTTCAGCTTGGTCA 840  
DB 2648 TGGGAGACATCGCCACTGSCATCAGCAGCAGATCCAGCTGCGACCTTCAGCTTGGTCA 2707  
QY 841 CCAAGACGGGAGGCTGTTCCTAGACATGGAGGTGCGAGCTCGGGCATTCGACCTGC 900  
DB 2708 CCAAGACGGGAGGCTGTTCCTAGACATGGAGGTGCGAGCTCGGGCATTCGACCTGC 2767  
QY 901 AGTGACACTGCGCCCTGATGACGCTTCGCTTGGAGCTGGAGGTCAAGCATGCGCAGC 960  
DB 2768 AGTGACACTGCGCCCTGATGACGCTTCGCTTGGAGCTGGAGGTCAAGCATGCGCAGC 2827  
QY 961 TGGAGAACAGGCCCTAA 977  
DB 2828 TGGAGAACAGGCCCTAA 2844

RESULT 13  
ACN44438  
ID ACN44438 standard; DNA; 73967 BP.  
XX AC  
AC ACN44438;  
XX

DT 18-NOV-2004 (first entry)

XX Human genomic sequence hCG27607.

DE Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.

XX WO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US0006235.

XX

PR 01-MAR-2002; 2002US-00087192.  
XX (SAGR-) SAGRES DISCOVERY.  
XX PI Morris DW;  
XX DR WPI; 2003-328604/31.  
XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
XX  
PS Claim 1; SEQ ID NO 886; Opp; English.  
XX  
CC The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA coding  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published  
XX  
SQ Sequence 73967 BP; 18655 A; 18476 C; 19039 G; 17797 T; 0 U; 0 Other;  
Query Match 37.0%; Score 973.4; DB 11; Length 73967;  
Best Local Similarity 84.4%; Pred. No. 5.1e-260;  
Matches 1487; Conservative 5; Mismatches 183; Indels 87; Gaps 35;  
QY 811 AGATCCAGCTGACGCTTTCAGTTGTGTCACCAAGACGGGACGCTTTCGCTACGACA 870  
DB 62243 AGATCCAGCTGACGCTTTCAGTTGTGTCACCAAGACGGGACGCTTTCGCTACGACA 62302  
QY 871 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGGCCCTTGATGGAGCTTCG 930  
DB 62303 TGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGGCCCTTGATGGAGCTTCG 62362  
QY 931 CCTCGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCAC 990  
DB 62363 CCTCGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCAC 62422  
QY 991 CGCCGGTCCACACTGCGGAAAGACGCTTTCCTGCTCGGTGACAGATGTGCCCTGAAA 1050  
DB 62423 CGCCGGTCCACACTGCGCGG-AAGCAGCTTCTCTGCTCGGTGACAGATGTGCCCTG-AA 62480  
QY 1051 ACACAGGCTCAGCCGTTCCAGGGGATYTGNCAGCCCCCGGCTCAGCAGNTGGGAACC 1110  
DB 62481 ACACAGGCTCAGCCGTTCCAGGGG---ATGTGCGAGCCCCCGGCTCAGCAGNTGGGAACC 62537  
QY 1111 AGGCGCTTCGNACGNACNAGAGTNGGGGCAAGNAGATGCTCCAGGATTCACAN 1170  
DB 62538 GGCCTCCAGCAGCAAG-----GTGGGGCAAGCAGATGCTCCAGGATTCACAC 62590  
QY 1171 CCTGAGCCNTGCCCCANCCCTGCTGAADAAAACAYTNCGCCACAGTGAAGAGACAGAA 1230  
DB 62591 C---TGAGCCCTGCCCCACGCTGCTGAGAAAAACACT--CGGCCACGTGAAGAGACA--- 62641  
QY 1231 GAGGATGNCAGGAGTTNNACCTYGGGAAACAAAACAGGAGATCTTTTCTGCCCTGC 1290  
DB 62642 GAGGAGGATGGCAGAGGTACCTCGGGAACAAACAGGATCTTCTCTGCCCTGCTCCAG- 62700  
QY 1291 TCCAGTNCAGGTTGGCCTGNACCCGCTTGGANTCAGTGCACATTTGTTGGCAGANCAAGG 1350  
DB 62701 -----TCGAGTTGGCTGA--CCCGCTTGAATCAGTGCACATTTGCTGGCAGA-CAGGG 62751  
QY 1351 GAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGGAGCCCTTTCGGCCCTCACCCTNC 1410  
DB 62752 GAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGGAGCCCTTTCGGCCCTCACCCT-C 62810



QY 1411 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGNGCCAAANCTCAGGNTTCAGTGCAGA 1470  
Db 62811 CAGGCTGCTGTG-AGAGTGTCAAGTGTGTAAAGGNGCCAAACTC---AGGTTCAAGTGCAGA 62866  
QY 1471 ACCAGGTCAGCAGTATGCCCGCCGNTAGTGTAAANNNGGGGCGCTCTNAAACCCCTTG 1530  
Db 62867 ACCAGGT-CAGCAGTATGCCCGCCGCTAGGTAA----GGGGGCGCTCTAAACCCCTTG 62921  
QY 1531 CTCTGGCCCTNACCTTNGCCAGCTCANCCCTTTTGGGTGTAGGGGAAAGAAATGCCCTGA 1590  
Db 62922 CTGSCCTCA---CCTGSCCAGCTCA-CCCCTTTTGGGTGTAGGGGAAAGAAATGCCCTGA 62977  
QY 1591 CCCTGGGAAGCTWCCCTGTGTAGATAACACACTTTTTCAGGTTGTGCAACACAGGTC 1650  
Db 62978 CCCTGGGAAGCT-CCCTGTGTAGATAACACACTTTTTCAGGTTGTGCAACACAGGTC 63036  
QY 1651 CTGAGTTGACCTCTGTGTTTACGCAAGGACCAAGAGGTGTAAAGTGAAGTGTCTTCA 1710  
Db 63037 CTGAGTTGACCTCTGTGTTTACGCAAGGACCAAGAGGTGTAAAGTGAAGTGTCTTCA 63096  
QY 1711 GTNCCCAGACATGTGCCCTTTGCTGTGCTTACCACTCTTCCCCAGAGCAGAGGCC 1770  
Db 63097 GT-CCCCAGACATGTGCCCTTTGCTGTGCTTACCACTCTTCCCCAGAGCAGAGGCC 63155  
QY 1771 CGAGCCCTTCAGGCCAGCACTGCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTA 1830  
Db 63156 CGAGCCCTTCAGGCCAGCACTGCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTA 63215  
QY 1831 AAGTGAAGGTGTATGAGGATATGCTGTAGAGCAAGTGTGGATGAGCATGATCAG 1890  
Db 63216 AAGTGAAGGTGTATGAGGATATGCTGTAGAGCAAGTGTGGATGAGCATGATCAG 63275  
QY 1891 TGTCTAAAGNAAAGCAGCAGAGAGAGCGYTCGGCGGCCCCAGNCCCACCTNATCAGTGT 1950  
Db 63276 TGTCTAA--GGAAAGCAGCAGAGAGAGCGTTCGGCGGCCCCAGNCCCACCTATCAGTGT-- 63331  
QY 1951 NCCAGGTGTCTNCCAGNAGCAGCACTTNCAGNATPANCACCTGACACTTNCACCC 2010  
Db 63332 -CCAGCGTGTCTGTTCC------CAGAGCAGCAGTCTCAGCATCAGACTCACC 63382  
QY 2011 TNGCCCTGCCCTTNGCCANGAGGTTACTCGGACAGGCACTTTCAGCTGTGAGACC 2070  
Db 63383 CTGCCCCTGCCCTTG-CCAGAGGATGTCGCG-ACGGCACTTTGCAC-TCTGATG-ACC 63437  
QY 2071 TCAAAGCACTTTTCATGGCTNCCCTCTTNNCGAGGNCAGGNCAGGNCAGTGCACACT 2130  
Db 63438 TCAAAGCACTTTTCATGGCTGCTCTTGGCAGGGCAGGGCAGGGCAGTGCACACT----- 63490  
QY 2131 GTAGGAGCATPANGCAANGCCAGGAGATGGGGTGAAGGAGNACAGTCTTTGAGCTGTCC 2190  
Db 63491 ---GTAGGAGCATAGCAAGCCAGGAGATGGGGTGA--GGGACACAGTCTTGGAGCTGCC 63545  
QY 2191 ANCATGATGATNCTTAACTCTTNNCCAGNATTTCTTAAGATAGNCCCTCC 2250  
Db 63546 A-CATGATGTGACT-CCTCAAACCTCTT---CCAGATTTCTTAAGATAGCACCCCTCC- 63599  
QY 2251 TTNCCCATTGCCCCAGCTAGCTCTTCTCCAGGGGAGCTANCTCAGGACTCACCTAG 2310  
Db 63600 -TTCCCCATTGCCCCAGCTAGCTCTTCTCCAGGGGAGCTA-CTCAGGACTCACCTAG 63657  
QY 2311 CATTAAATCAGCTGTGNAATCGTCAGGGGGTGTCTGTAGCTCAACCTCTCTGGGGCAGG 2370  
Db 63658 CATTAAATCAGCTGTG-AATGCTCAGGGGGTGTCTGTAGCTCAACCTCTCTGGGGCAGG 63716  
QY 2371 GGAGCCGAGACTCCGTGGGGAAGAGTCAATTCACATCTTGGCAAGACAGCCTTTNGTC 2430  
Db 63717 GGAGCCGAGACTCCGTGGGGAAGAGTCAATTCACATCTTGGCAAGACAGCCTTT-GTC 63775  
QY 2431 CAGCTGTCCATTTAGTGTGACTGCTCCGGGAGAGAGCCCGCCCGCCAGACATAA 2490  
Db 63776 CAGCTGTCCATTTAGTGTGACTGCTCCGGGAGAGAGCCCGCCCGCCAGACATAA 63835

QY 2491 AGAACTGCAGCCTTGGTACTGCAGAGTCTGGTGTGTAGAGAACTCTTTGTGAAGCAATAA 2550  
Db 63836 AGAACTGCAGCCTTGGTACTGCAGAGTCTGGTGTGTAGAGAACTCTTTGTGAAGCAATAA 63895  
QY 2551 GTTTGGGGTGATGACAAATGTT 2572  
Db 63896 GTTTGGGGTGATGACAAATGTT 63917  
RESULT 14  
AAH99262  
ID AAH99262 standard; cDNA; 1385 BP.  
XX  
AC AAH99262;  
XX  
DT 16-OCT-2001 (first entry)  
XX  
DE Human protein encoding cDNA sequence SEQ ID NO:97.  
XX  
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
KW antibacterial; endocrine; cardiac; antidiabetic; central nervous system; virucide;  
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
KW antiaggregant; haemostatic; vulnery; antidiabetic; osteopathic; eczema;  
KW dermatological; antiallergic; antidiabetic; antidiabetic; cytostatic;  
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
KW immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;  
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
KW neurological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200153455-A2.  
XX  
PD 26-JUL-2001.  
XX  
PF 22-DEC-2000; 2000WO-US035017.  
XX  
PR 23-DEC-1999; 99US-00471275.  
PR 21-JAN-2000; 2000US-00488725.  
PR 25-APR-2000; 2000US-00552317.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
PI WPI; 2001-457603/49.  
XX  
DR P-PSDB; AAM25321.  
XX  
XX Isolated human polynucleotides encoding oligopeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
PS Claim 1; Page 342-343; 1217pp; English.  
XX  
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antirheumatic; antidiabetic; immunosuppressive; antibacterial; endocrine; cardiac; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antidiabetic; haemostatic; vulnery; antidiabetic; cytostatic; dermatological; antiallergic; antidiabetic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antitense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

CC	neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC	infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC	anemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC	osteoporosis, severe combined immunodeficiency, eczema, allergic
CC	rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC	Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC	neurological disorders
XX	
SQ	Sequence 1385 BP; 320 A; 410 C; 378 G; 277 T; 0 U; 0 Other;
	Query Match 28.7%; Score 754.2; DB 4; Length 1385;
	Best Local Similarity 83.5%; Pred. No. 1.2e-199;
	Matches 1205; Conservative 4; Mismatches 159; Indels 75; Gaps 31;
QY	1137 GGGCAAGCAGGATGCTCCAGGATTTTCACANCTTGAGCCNTGCCANCCCTCGT 1196
DB	15 GGGGCAAGCAGATGCTCCAGGATTTTCACACC---TGAGCCCTGCCCAACCTGCTG 71
QY	1197 AADAAACATYNCGCCACGTGAAGACACAGAGGAGGATGNCAGAGTTNNACCTYGG 1256
DB	72 AGAAAAACAT--CGGCCACGTGAAGAGACA---GAGGAGGATGGCAGGAGTTACCTCGG 125
QY	1257 GGAACAAAAACAGGATCTTTNTTCTGCCCTGCTCCAGTNCAGATGGCTGNACCCGC 1316
DB	126 GAACAACACAGGATCTTCTTGCCCTGCTCCAG-----TCGAGTTGGGCTG-ACCCGC 177
QY	1317 TTGGANTCAGTGACCAATTTGTGGCAGANCAGGGGAGCAGCAGCTTCACAGCTGGGTGAGA 1376
DB	178 TTGGA-TCAGTGACCAATTTGCTGGCAGA-CAGGGGAGACAGCTTCAGAGCTGGGTGAGA 235
QY	1377 AGGGGTGGCGAGCCCTTCGGCCCTCACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTG 1436
DB	236 AGGGGTGGCGAGCCCTCGGCCCTCACCT-CCAGGCTGCTGTG-AGAGTGTCAAGTG 293
QY	1437 TGTAAAGGNCACCAANCTCAGGNTTCAGTGCAGAACACAGTNCAGCAGGTATGCCCGCCC 1496
DB	294 TGTAAAGGCCCAAACTC---AGGTTTCAGTGCAGAACACAGT-CAGCAGGTATGCCCGCCC 349
QY	1497 GNTAGGTTAANNGGGGGCCCTCTNAAACCCCTTGCTNGCCCTNCACCTNGGCCAGCTCA 1556
DB	350 GTACGTTAA----GGGGGCCCTCTAAACCCCTTGCTTGCCCTGAGCTCA---CTTGGCAGCTCA 402
QY	1557 NCCTCTTTTGGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAAGGCTTWCCTCGTGAAT 1616
DB	403 -CCCTTTTGGGTGTAGGGGAAAAGATGCTGACCCCTGGGAAAGGCT-CCCTGGTGAAT 460
QY	1617 ACACCACACTTTTCAGTGTGTGGAACACACAGGTCTGAGTTGACCTGTGTTGAGCCAAAG 1676
DB	461 ACACCACACTTTTCAGTGTGTGGAACACACAGGTCTGAGTTGACCTGTGTTGAGCCAAAG 520
QY	1677 GACCAAGAAGGTGTGTAAAGTGAAGTGGTTCCTCAGTNCACCCAGACATGTGCCCTTTGCT 1736
DB	521 GACCAAGAAGGTGTGTAAAGTGAAGTGGTTCCTCAGT-CCCCAGACATGTGCCCTTTGCT 579
QY	1737 GCTGGCTTACCACCTTTCCCCAGACAGCAGGCCCCCGAGCCCTTCAGGCCCCAGCACTGCC 1796
DB	580 GCTGGCTTACCACCTTTCCCCAGACAGCAGGCCCCCGAGCCCTTCAGGCCCCAGCACTGCC 639
QY	1797 CCAGACTCGCTGGGACCTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGATCAGGATATGC 1856
DB	640 CCAGACTCGCTGGGACCTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGATCAGGATATGC 699
QY	1857 CTCACAGGAACACTGTGTGATGGACATGATCAGTCTNAAAGNAAAGCAGCAGAGAG 1916
DB	700 CTCACAGGAACACTGTGTGATGGACATGATCAGTCTAA--GGAAAGCAGCAGAGAGAG 757
QY	1917 ACGYTCGGCGCCCCCAGNCCCCTNATCAGTGTNCCAGCGCTGCTNGGTTNCCCCAGNAG 1976
DB	758 ACGCTCGGCGCCCCCAGCCCCCATATCAGTGT---CCAGGCTGCTGGTTCCT-----806
QY	1977 CACAGCTNCAGNCAATCANCACTGACACTNCACCTCTNGCCCTGCCCCCTNGGCCANGAGGT 2036
DB	807 CACAGCAGAGCTCAGCATACACTGACACTCACTGACCTGCCCTGCCCTCTGG--CCAGAGGGT 864

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX  
PS Claim 1; SEQ ID NO 15295; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 2760 BP; 486 A; 840 C; 810 G; 624 T; 0 U; 0 Other;

Query Match 26.3%; Score 691.8; DB 5; Length 2760;  
Best Local Similarity 98.3%; Pred. No. 4.6e-182;  
Matches 699; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY	267	AAGCTCAGGCTCTCTCCACAGAGAGCCCGGAGCAAAAGTCTCTCCCTTGACT	326
DB	711	AAGGTGAACCGGGCAGCTACAGCAAGAGAGCCCGGAGCAAAAGTCTCTCCCTTGACT	652
QY	327	TTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTTACTCTGTCTCCCTGGAGCCAGCC	386
DB	651	TTGAGCAAGGAGGAGTCTGGGATGTGGAAACCTTACTCTGTCTCCCTGGAGCCAGCC	592
QY	387	CTGTCCAGAAACCCAGCTCACAGAGCGGAAACCGTCCCGAGCAGGAACCTGCAG	446
DB	591	CCTGCCAGAAACCCAGCTCACAGAGCGGAAACCGTCCCGAGCAGGAACCTGCAG	532
QY	447	CAGCTGAAATAGAAATTTCTCTCAACAGCGTGTCCAGCCATTTCTCTGGAGGAGCAG	506
DB	531	CAGCTGAAATAGAAATTTCTCTCAACAGCGTGTCCAGCCATTTCTCTGGAGGAGCAG	472
QY	507	GAGCAAAATCTCTGTGCTCAGCATCGAGCTCTCCCTGTGGATGACAGTGAGAA	566
DB	471	GAGCAAAATCTCTGTGCTCAGCATCGAGCTCTCCCTGTGGATGACAGTGAGAA	412
QY	567	AACCCATCAAAAGGCTCTCAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCCTGG	626
DB	411	AACCCATCAAAAGGCTCTCAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCCTGG	352
QY	627	AGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGAAACATGGTGTGGCCCGGGGGCGGCC	686
DB	351	AGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGAAACATGGTGTGGCCCGGGGGCGGCC	292
QY	687	ACCGACACCCCAAGCTATTTCAATGGTGTCAAAAGTCCAAATACAGTCTCTTAATGTGAA	746
DB	291	ACCGACACCCCAAGCTATTTCAATGGTGTCAAAAGTCCAAATACAGTCTCTTAATGTGAA	232
QY	747	CACCTGCACATCCGGGAGTTCCACCGGGTCAAAAGTGGAGACATCGCCACTGCATCAGC	806
DB	231	CACCTGCACATCCGGGAGTTCCACCGGGTCAAAAGTGGAGACATCGCCACTGCATCAGC	172
QY	807	AGCCAGATCCAGCTGAGGCTTTCAGCTTGGTACCAAAAGCGGAGCCTGTTCGCTAC	866

Db	171	AGCCAGATCCCAGCTCGAGCCTTTCAGCTTGGTCAACCAAGACGGGAGCCTGTTCGTAC	112
QY	867	GACATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCTGTATGGCAGC	926
Db	111	GACATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCTGTATGGCAGC	52
QY	927	TTGCGCTTGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA	977
Db	51	TTGCGCTTGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA	1

Search completed: June 10, 2005, 06:10:27  
Job time : 914.226 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:18:24 ; Search time 7480.59-Seconds  
(without alignments)  
17042.190 Million cell updates/sec

Title: US-09-155-676B-3  
Perfect score: 2631  
Sequence: 1 cccctctcacagccagcc.....acgtgcacctcgaggggg 2631

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2553.8	97.1	2631	6	A66647
2	1851.2	70.4	4596	6	A66650
3	1850.2	70.3	4596	6	AX429240
4	1850.2	70.3	4596	6	AX774978
5	1850.2	70.3	4596	9	HSNIK
6	1788.8	68.0	4452	9	BC035576
7	1781.8	67.7	4584	6	CQ723104
8	1771.6	67.3	3152	6	CQ845924
9	1771.6	67.3	3152	9	AK131438
10	1133.4	43.1	3156	6	AR062290
11	1133.4	43.1	3156	6	AR063065
12	1133.4	43.1	3156	6	AR068515
13	1133.4	43.1	3156	6	BD062401
14	977	37.1	2844	6	AB429673
15	973.4	37.0	48032	9	AC003963
16	973.4	37.0	96875	9	AC008105
17	973.4	37.0	100836	2	AC087298
18	971	36.9	142326	2	AC024047
19	672.8	25.6	2829	10	AF143094

c	20	349.4	13.3	2475	6	AX834630	AX834630 Sequence
c	21	349.4	13.3	2475	9	AK097219	AK097219 Homo sapi
c	22	283	10.8	337	11	G27543	G27543 human STS S
c	23	275.8	10.5	162192	2	AC150796	AC150796 Bos tauru
c	24	266.6	10.1	70748	2	AC137575	AC137575 Homo sapi
c	25	238.4	9.1	698	11	BV166476	BV166476 MAP3K14_9
c	26	205.8	7.8	476	6	AX778201	AX778201 Sequence
c	27	188.8	7.2	213625	2	AC120950	AC120950 Rattus no
c	28	188.8	7.2	223438	10	AL662804	AL662804 Mouse DNA
c	29	188.8	7.2	239464	2	AC103079	AC103079 Rattus no
c	30	138	5.2	70748	2	AC137575	AC137575 Homo sapi
c	31	100.6	3.8	101	11	BV184269	BV184269 sqm14402
c	32	94.2	3.6	374	10	AY091549S1	AY091549 Mus muscu
c	33	91	3.5	907	9	BC031942	BC031942 Homo sapi
c	34	88.4	3.4	455	6	CQ490185	CQ490185 Sequence
c	35	88.4	3.4	455	6	CQ496042	CQ496042 Sequence
c	36	88.4	3.4	476	6	CQ480458	CQ480458 Sequence
c	37	88.4	3.4	504	6	CQ501629	CQ501629 Sequence
c	38	88.4	3.4	504	6	CQ510553	CQ510553 Sequence
c	39	73.8	2.8	463	6	CQ471290	CQ471290 Sequence
c	40	66	2.5	125020	9	AF429315	AF429315 Homo sapi
c	41	63.8	2.4	125020	9	AF429315	AF429315 Homo sapi
c	42	60.4	2.3	111803	2	AC136172	AC136172 Rattus no
c	43	58.2	2.2	2000	6	AX655393	AX655393 Sequence
c	44	53.2	2.0	1253	6	A66648	A66648 Sequence 4
c	45	52.8	2.0	65	6	CQ556101	CQ556101 Sequence

ALIGNMENTS

RESULT 1	LOCUS	A66647	2631 bp	DNA	linear	PAT 29-MAR-1999
DEFINITION	Sequence 3 from Patent WO9737016.					
ACCESSION	A66647					
VERSION	A66647.1	GI:4538139				
KEYWORDS	unidentified					
SOURCE	unidentified					
ORGANISM	unclassified.					
REFERENCE	1 (bases 1 to 2631)					
AUTHORS	Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.					
TITLE	MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR PREPARATION AND USE					
JOURNAL	Patent: WO 9737016-A 3 09-OCT-1997;					
COMMENT	YEDA RES & DEV (IL)					
FEATURES	Other publication AU 2175597 19971022.					
	Location/Qualifiers					
	source					
	1..2631					
	/organism="unidentified"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32644"					
ORIGIN						
	Query Match	97.1%	Score 2553.8;	DB 6;	Length 2631;	
	Best Local Similarity	100.0%;	Pred. No. 0;			
	Matches 2631;	Conservative	0;	Mismatches	0;	Gaps 0;
Qy	1	CCCTCTTCACAGCCCCAGGCCATCCAGAGGGGGCTGAGGAAAGAGCCCCATCCACCGCGTGT	60			
Db	1	CCCTCTTCACAGCCCCAGGCCATCCAGAGGGGGCTGAGGAAAGAGCCCCATCCACCGCGTGT	60			
Qy	61	CTGCAGGGAGCTGGGAGGAGGTGAACCGGGGCACTACAGCAAGTGGAGGCTCTGAAGA	120			
Db	61	CTGCAGGGAGCTGGGAGGAGGTGAACCGGGGCACTACAGCAAGTGGAGGCTCTGAAGA	120			
Qy	121	GCCCTTCGAGGGAGAGATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC	180			
Db	121	GCCCTTCGAGGGAGAGATATAAGAAACCAAGACATCCACCGCCCAATTCAAGCCCAATTACC	180			
Qy	181	ACCAGACCTTCATGCTCCAGCCAGGAGAGCTTTTCGCCAAGGGGCCCGCGGCCAG	240			
Db	181	ACCAGACCTTCATGCTCCAGCCAGGAGAGCTTTTCGCCAAGGGGCCCGCGGCCAG	240			

181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTGGCCMAAGGCCCCAGGGCCCCCGGCCAG 240 Db  
241 CTGAGGAGACAACAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCCACACGAGAGCCCCAG 300 Qy  
241 CTGAGGAGACAACAGGCGAGAGCCCTTAAGCTCCAGCTCTCTCCACACGAGAGCCCCAG 300 Db  
301 AGCCAAAACAAGTCTCTCTCCCTGACTTTGAGCAGAGGAGGTCTGGGATGTGGAAACCT 360 Qy  
301 AGCCAAAACAAGTCTCTCTCCCTGACTTTGAGCAGAGGAGGTCTGGGATGTGGAAACCT 360 Db  
361 TACCTCTGTCTCTCCCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAG 420 Qy  
361 TACCTCTGTCTCTCTCCCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAG 420 Db  
421 CAACCGTCCCGAGCAGGAATCTGACGAGCTGGAATAAGATTAATTCCTCAACAGCCTGT 480 Qy  
421 CAACCGTCCCGAGCAGGAATCTGACGAGCTGGAATAAGATTAATTCCTCAACAGCCTGT 480 Db  
481 CCGAGCATTTTCTCTGGAGGAGGAGGAAATTTCTCTCGTGCCTCAGCATCGACAGCC 540 Qy  
481 CCGAGCATTTTCTCTGGAGGAGGAGGAAATTTCTCTCGTGCCTCAGCATCGACAGCC 540 Db  
541 TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAAAGCTCGCGGACA 600 Qy  
541 TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAAAGCTCGCGGACA 600 Db  
601 CCTGAGCTCAGCGTACACTCTGTGAGCAGCCAGCCGAGCTCGAAGCTCCAGCTGGA 660 Qy  
601 CCTGAGCTCAGCGTACACTCTGTGAGCAGCCAGCCGAGCTCGAAGCTCCAGCTGGA 660 Db  
661 ACATGTGTGCGCCCGGGGCGCCACCGACACCCCAAGCTATTCAATGGTGTGAAAG 720 Qy  
661 ACATGTGTGCGCCCGGGGCGCCACCGACACCCCAAGCTATTCAATGGTGTGAAAG 720 Db  
721 TCCAAATACAGTCTCTTAATGTGAAACACTGACATCGGAGTTTCCACCGGGTCAAG 780 Qy  
721 TCCAAATACAGTCTCTTAATGTGAAACACTGACATCGGAGTTTCCACCGGGTCAAG 780 Db  
781 TGGGAGACATCGGCATCGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840 Qy  
781 TGGGAGACATCGGCATCGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840 Db  
841 CCAAAGCGGCGAGCTGTTGCTGACGATGAGGTGCCAGACTCGGGCATCGACCTGC 900 Qy  
841 CCAAAGCGGCGAGCTGTTGCTGACGATGAGGTGCCAGACTCGGGCATCGACCTGC 900 Db  
901 AGTGCACTGGGCCCTGTATGGCAGCTTGCCTTGGAGCTGGAGGTCAAGCATGGCCAG 960 Qy  
901 AGTGCACTGGGCCCTGTATGGCAGCTTGCCTTGGAGCTGGAGGTCAAGCATGGCCAG 960 Db  
961 TGGAGAACGCGCTTACCTGCGCTCCACCGCGGCTCCACACTGCGCGAAAGCAGCCT 1020 Qy  
961 TGGAGAACGCGCTTACCTGCGCTCCACCGCGGCTCCACACTGCGCGAAAGCAGCCT 1020 Db  
1021 TCCTGCTCGGTGCAAGTGTGCTTGAACACAGGCTCAGCGTTCAGGGGATYTG 1080 Qy  
1021 TCCTGCTCGGTGCAAGTGTGCTTGAACACAGGCTCAGCGTTCAGGGGATYTG 1080 Db  
1081 NCCAGCCCCCGGCTCARGAGNTGGAAACAGGGCTTCGNCAGNAGNAGNTGGGG 1140 Qy  
1081 NCCAGCCCCCGGCTCARGAGNTGGAAACAGGGCTTCGNCAGNAGNAGNTGGGG 1140 Db  
1141 CAAGNAGATGCTTCCAGGATTTACANCCTGAGCCNTGCCCCANCCCTGCTGAADA 1200 Qy  
1141 CAAGNAGATGCTTCCAGGATTTACANCCTGAGCCNTGCCCCANCCCTGCTGAADA 1200 Db  
1201 AAAAYTNCCGCCAGTGAAGAGACAGAGAGATGNCAGGATTTNNAACCTYGGGAA 1260 Qy  
1201 AAAAYTNCCGCCAGTGAAGAGACAGAGAGATGNCAGGATTTNNAACCTYGGGAA 1260 Db  
1261 ACMAACAGGATCTTTNTTCTGCTGCTTCCAGTTCGAGTTGGCTTGCCTGNAACCCGCTTG 1320 Qy  
1261 ACMAACAGGATCTTTNTTCTGCTGCTTCCAGTTCGAGTTGGCTTGCCTGNAACCCGCTTG 1320 Db

1321 ANTCACTGACCATTTGTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 1380 Qy  
1321 ANTCACTGACCATTTGTTGGCAGANCAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 1380 Db  
1381 GTGGGCGAGCCCTTCGGCCCCCTCACCCCTNCCAGGCTGTGTGNAGAGTGTCAAGTGTGTA 1440 Qy  
1381 GTGGGCGAGCCCTTCGGCCCCCTCACCCCTNCCAGGCTGTGTGNAGAGTGTCAAGTGTGTA 1440 Db  
1441 AGGNCCTTCAAACTCAGGNTTCAGTCAGAACCCAGGTCAGAGGTATGCCCGCCGNTA 1500 Qy  
1441 AGGNCCTTCAAACTCAGGNTTCAGTCAGAACCCAGGTCAGAGGTATGCCCGCCGNTA 1500 Db  
1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCCTNNGGCTNCACTNNGGCGAGCTCANCCC 1560 Qy  
1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCCTNNGGCTNCACTNNGGCGAGCTCANCCC 1560 Db  
1561 CTTTGGGTGTAGGGGAAAAGATGCTGACCTCGGGAAGGCTTWCCTTGTAGAAATACAC 1620 Qy  
1561 CTTTGGGTGTAGGGGAAAAGATGCTGACCTCGGGAAGGCTTWCCTTGTAGAAATACAC 1620 Db  
1621 CACACTTTTCAGGTGTGTGCAACACAGGTCCTCAGTTGACCTCTGGTTCAGCCAAAGACC 1680 Qy  
1621 CACACTTTTCAGGTGTGTGCAACACAGGTCCTCAGTTGACCTCTGGTTCAGCCAAAGACC 1680 Db  
1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCAGACATGTGCCCCCTTTGCTGCTG 1740 Qy  
1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCAGACATGTGCCCCCTTTGCTGCTG 1740 Db  
1741 GCTACACTCTTCCCCAGAGCAGCCGCGGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 1800 Qy  
1741 GCTACACTCTTCCCCAGAGCAGCCGCGGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 1800 Db  
1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTA 1860 Qy  
1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGAGGATATGCTGTA 1860 Db  
1861 CAGAAACAGTCTGTGATGACATGATCAGTGTCTNAGGNAAGCAGCAGAGAGACGY 1920 Qy  
1861 CAGAAACAGTCTGTGATGACATGATCAGTGTCTNAGGNAAGCAGCAGAGAGACGY 1920 Db  
1921 TCCGGGCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTCTNNGTTNCCCAGNAGACA 1980 Qy  
1921 TCCGGGCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTCTNNGTTNCCCAGNAGACA 1980 Db  
1981 GCTNAGNATCANCACTGACACTNCACTTNGCCCTGCCCCCTNGGCCANGAGGTACTG 2040 Qy  
1981 GCTNAGNATCANCACTGACACTNCACTTNGCCCTGCCCCCTNGGCCANGAGGTACTG 2040 Db  
2041 CCGNACGGCACTTTGCACTCTGATGNACCTCAAAGCACTTTTATGGCTNGCCCTCTNNG 2100 Qy  
2041 CCGNACGGCACTTTGCACTCTGATGNACCTCAAAGCACTTTTATGGCTNGCCCTCTNNG 2100 Db  
2101 GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAANGCCAGGAGATGG 2160 Qy  
2101 GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAANGCCAGGAGATGG 2160 Db  
2161 GGTGNAAGGNCACAGTCTTGTGAGCTGTCCANCATGTGACTNCTCAAACTCTTN 2220 Qy  
2161 GGTGNAAGGNCACAGTCTTGTGAGCTGTCCANCATGTGACTNCTCAAACTCTTN 2220 Db  
2221 NCCAGNATTTCTTAAGAAATAGCAGCCCTTNCCTTNGCCCTTCCAGCTTAGCCTCTTCT 2280 Qy  
2221 NCCAGNATTTCTTAAGAAATAGCAGCCCTTNCCTTNGCCCTTCCAGCTTAGCCTCTTCT 2280 Db  
2281 CCCAGGGAGCTTANCTCAGGACTCAGTAGCAATTAATCAGCTGTGNAATCGTCAGGGGG 2340 Qy  
2281 CCCAGGGAGCTTANCTCAGGACTCAGTAGCAATTAATCAGCTGTGNAATCGTCAGGGGG 2340 Db  
2341 TGTCTGTAGCTTCAACCTCTGGGGCAGGAGCCCGAGACTCCGCTGGGAGAGCTCAT 2400 Qy  
2341 TGTCTGTAGCTTCAACCTCTGGGGCAGGAGCCCGAGACTCCGCTGGGAGAGCTCAT 2400 Db

Qy 2401 TCCACATCTTGGCAAGACAGCCTTTNGTTCAGCTGTCCACATTTAGTTCAGACTGCTCCC 2460  
Db 2401 TCCACATCTTGGCAAGACAGCCTTTNGTTCAGCTGTCCACATTTAGTTCAGACTGCTCCC 2460  
Qy 2461 GGGAGAGAGCCCGGCCCCCAGACACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520  
Db 2461 GGGAGAGAGCCCGGCCCCCAGACACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520  
Qy 2521 GGTGTAGAGACTCTTTTGTAAAGCAATAAAGTTTGGGTGTATGACAAATGTTAAAAAAG 2580  
Db 2521 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGTATGACAAATGTTAAAAAAG 2580  
Qy 2581 GCCTTCGTGGCCTCGAATCAAGCTTATCGATACCGTCGACCTCGAGGGGG 2631  
Db 2581 GCCTTCGTGGCCTCGAATCAAGCTTATCGATACCGTCGACCTCGAGGGGG 2631  
RESULT 2  
LOCUS A66650 A66650 4596 bp DNA linear PAT 29-MAR-1999  
DEFINITION Sequence 6 from Patent WO9737016.  
ACCESSION A66650  
VERSION A66650.1 GI:4538141  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 4596)  
AUTHORS Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.  
TITLE MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR  
PREPARATION AND USE  
JOURNAL Patent: WO 9737016-A 6 09-OCT-1997;  
COMMENT YEDA RES & DEV (IL)  
Other publication AU 2175597 19971022.  
FEATURES  
Location/Qualifiers  
source 1..4596  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 70.4%; Score 1851.2; DB 6; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;  
Qy 1 CCCCTCTCACAGCCAGGCCATCCAAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 60  
Db 2099 CCCCTCTCACAGCCAGGCCATCCAAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 2158  
Qy 61 CTGACGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
Db 2159 CTGACGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2218  
Qy 121 GCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180  
Db 2219 GCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 2278  
Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCAAGGGCCCGAGGGCCCGGCCAG 240  
Db 2279 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCAAGGGCCCGAGGGCCCGGCCAG 2338  
Qy 241 CTGAGGAGACAAAGGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGAGCCCGCCAG 300  
Db 2339 CTGAGGAGACAAAGGAGAGCCCTTAAGTCTCAGCCTCTCTCCACAGAGAGCCCGCCAG 2398  
Qy 301 AGCCAAACAAGTCTCTCCCTTGAATTTGACAAAGAGAGTCTGGGATGTGGAAACCT 360  
Db 2399 AGCCAAACAAGTCTCTCCCTTGAATTTGACAAAGAGAGTCTGGGATGTGGAAACCT 2458  
Qy 361 TACCTGTCTCTCCCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAAG 420  
Db 2459 TACCTGTCTCTCCCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAAG 2518

Qy 421 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCTCTCAACAGCCTGT 480  
Db 2519 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCTCTCAACAGCCTGT 2578  
Qy 481 CCCAGCATTCTTCTCGAGGAGCAGGAGAAATTTCTCTCGTGTCTCAGCATCGACAGCC 540  
Db 2579 CCCAGCATTCTTCTCGAGGAGCAGGAGAAATTTCTCTCGTGTCTCAGCATCGACAGCC 2638  
Qy 541 TCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCTCAAGCTTCGCGGGACA 600  
Db 2639 TCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCTCAAGCTTCGCGGGACA 2698  
Qy 601 CCCTGAGCTCAGGCGTACACTCTCGAGAGCAGCAGGCGGAGCTCGAAGCTTCAGCTCGA 660  
Db 2699 CCCTGAGCTCAGGCGTACACTCTCGAGAGCAGCAGGCGGAGCTCGAAGCTTCAGCTCGA 2758  
Qy 661 ACATGTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCATGCTGTAAG 720  
Db 2759 ACATGTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCATGCTGTAAG 2818  
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCCGGAGTTCCACCGGCTCAAAG 780  
Db 2819 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCCGGAGTTCCACCGGCTCAAAG 2878  
Qy 781 TGGGAGACATCGCCACTTGGCATCAGAGCCAGATCCAGCTGCAGCTTCAAGCTTGTGCA 840  
Db 2879 TGGGAGACATCGCCACTTGGCATCAGAGCCAGATCCAGCTGCAGCTTCAAGCTTGTGCA 2938  
Qy 841 CCAAAGACGGGACGCTGTTTCGCTACACATGAGAGTGCCAGACTCGGGCATCGACCTGC 900  
Db 2939 CCAAAGACGGGACGCTGTTTCGCTACACATGAGAGTGCCAGACTCGGGCATCGACCTGC 2998  
Qy 901 AGTGACACTGGCCCTGTATGGCAGCTTCGCTTGGAGCTTGGAGGTCAAGCATGGCCAGC 960  
Db 2999 AGTGACACTGGCCCTGTATGGCAGCTTCGCTTGGAGCTTGGAGGTCAAGCATGGCCAGC 3058  
Qy 961 TGGAGAACAGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGCGGAAAGAGAGCT 1020  
Db 3059 TGGAGAACAGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGCGGAAAGAGAGCT 3118  
Qy 1021 TCCTGTCTGGTGACAGATGCTGCCCTGAAACACAGCTCAGCCGTTCCACAGGGGATGTG 1080  
Db 3119 TCCTGTCTGGTGACAGATGCTGCCCTGAAACACAGCTCAGCCGTTCCACAGGGGATGT-- 3176  
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCGCTTCGNACAGCNAAAGTNGGGGG 1140  
Db 3177 GCCAGCCCCCGGCTCA--CAGTGGGAACACAGGCGCTCG-----CAGCAGCAAGTGGGG 3229  
Qy 1141 CAAGCNAAGATGCTCCAGGATTTTCAANCCTGAGCCCTGCCCCANCCCTGCTGTAADA 1200  
Db 3230 GCAAGCAGAAATGCTCCAGGATTTTCAACTGAGCCCTGCCCCAC-----CCTGCTGA 3283  
Qy 1201 AAACATYNNCCGACGTTGAAGAGACAGAGGAGTGCNAGAGTNNACCTYGGGAA 1260  
Db 3284 AAAACATYNNCCGACGTTGAAGAGACAGAGGAGTGCNAGAGTNNACCTYGGGAA 3339  
Qy 1261 AAAAAACAGGAGTCTTTNTTCTGCCCCCTGCTCCAGTTCGAGTTGGCTTGCCTGNAACCCGCTGG 1320  
Db 3340 AAAAAACAGGAGTCTTTT-TTCTGCCCCCTGCTCCAGT--CGAGTTGGCTCGA--CCCGCTTG 3395  
Qy 1321 ANTCAGTGACCAATTTGTTCAGANACAGGGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 1380  
Db 3396 GATCAGTGACCAATTTGTTCGAGA-CAGGGAGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 3454  
Qy 1381 GTGGGCGAGCCCTTCGGCCCTCACCTNCCAGGCTCTGTGNAGAGTGCAGGTGTGA 1440  
Db 3455 GTGGGCGAGCCCTTCGGCCCTCACCTNCCAGGCTCTGTGNAGAGTGCAGGTGTGA 3512  
Qy 1441 AGGNCCTCAAACTCAGGNTTCAGTGCAGAACACAGTTCNACAGCAGTATGCCCGCCCGNTA 1500  
Db 3513 AGGNCCTCAAACTC---AGGTTCAGTGCAGAACACAGT--CAGCAGTATGCCCGCCCGTAG 3568  
Qy 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCCTNNGGCTTNCACCTNNGGCCAGCTCANCCC 1560





Db 2760 ACATGGTCTGGCCCGGGGGCCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2819  
Qy 721 TCCAAATAAGTCTCTTAATGGTGAACACCTGCAATCCGGGAGTTCCACCGGTCAAAG 780  
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGGAGTTCCACCGGTCAAAG 2879  
Qy 781 TGGGAGACATGCCACTGGCATCAGCAGCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 840  
Db 2880 TGGGAGACATGCCACTGGCATCAGCAGCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 2939  
Qy 841 CCAAAGACGGCAGCTCTTCGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 900  
Db 2940 CCAAAGACGGCAGCTCTTCGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 2999  
Qy 901 AGTCACACTGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTCACACTGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3059  
Qy 961 TGGGAAACAGGCCCTTAACCTGCTCCATCCGCGCGGTCCACACTGCCGGAAAGCAGCT 1020  
Db 3060 TGGGAAACAGGCCCTTAACCTGCTCCATCCGCGCGGTCCACACTGCCGGAAAGCAGCT 3119  
Qy 1021 TCCTGCTCGGTGACGATGCTGCTGAAACACAGAGCTCAGCGCTTCCAGGGGATYTG 1080  
Db 3120 TCCTGCTCGGTGACGATGCTGCTGAAACACAGAGCTCAGCGCTTCCAGGGGATYTG 3177  
Qy 1081 NCCAGCCCCCGGTCTACAGNTGGAAACCAAGGCTCTGNCAGNAGNAGTNGGGG 1140  
Db 3178 GCCAGCCCCCGGTCTA--CAGTGGAAACAGGGCTCG-----CAGCAGCAAGGTGGG 3230  
Qy 1141 CAAGCAGAAATGCTCCAGATTTTCAACNCTGAGCCGNTGCCANCCCTGCTGAADA 1200  
Db 3231 GCAAGCAGAAATGCTCCAGATTTTCAACNCTGAGCCGNTGCCANCCCTGCTGAADA 3284  
Qy 1201 AAACAYTNCGCCACGTGAAGACAGAGAGAGATGNCAGAGATTNNACCTVGGGAA 1260  
Db 3285 AAAAATCCGCCACGTGAAGACAGAGAGAGATGNCAGAGATTNNACCTVGGGAA 3340  
Qy 1261 AAAAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGATTGGCTGNAACCCGCTTG 1320  
Db 3341 AAAAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGATTGGCTGNAACCCGCTTG 3396  
Qy 1321 ANTCACTGACATTTGTCAGANCAAGGAGAGAGCTTCCAGCTGGGTGAGAAGG 1380  
Db 3397 GATCAGTGACATTTGTCAGAGAGAGAGCTTCCAGCTGGGTGAGAAGG 3455  
Qy 1381 GTGGCGAGCCCTTCGGCCCCCTCACCTNCAGAGCTGCTGAGAGTCAAGTGTGA 1440  
Db 3456 GTGGCGAGCCCTTCGGCCCCCTCACCTNCAGAGCTGCTGAGAGTCAAGTGTGA 3513  
Qy 1441 AGGNCCTCAAACTCAGNTTTCAGTGAGAACACAGGTCNACAGGTATGCCGCCCGNTA 1500  
Db 3514 AGGNCCTCAAACTCAGNTTTCAGTGAGAACACAGGTCNACAGGTATGCCGCCCGNTA 3569  
Qy 1501 GGTAAANNNGGGGCCCTTNAACCCCTTGCTTNGGCTNCACTNGGCCAGCTCANCCC 1560  
Db 3570 GTTAA----GGGGCCCTCTAAACCCCTTGCTTNGGCTNCACTNGGCCAGCTCANCCC 3621  
Qy 1561 CTTTTGGGTAGGGGAAAGAAATGCCTGAGCCCTGGGAGGCTWCCCTGGTAGAATACAC 1620  
Db 3622 CTTTTGGGTAGGGGAAAGAAATGCCTGAGCCCTGGGAGGCT--CCCTGGTAGAATACAC 3680  
Qy 1621 CACACTTTTCAAGTTTGGAAACAGAGCTGAGAGTTCAGCTCTGGTTTTCAGCCAGGACC 1680  
Db 3681 CACACTTTTCAAGTTTGGAAACAGAGCTGAGAGTTCAGCTCTGGTTTTCAGCCAGGACC 3740  
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCCAGACATGTCGCCCTTTGCTGCTG 1740  
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCCAGACATGTCGCCCTTTGCTGCTG 3799  
Qy 1741 GCTACCACTCTTCCCAAGAGCAGAGCCGCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800  
Db 3800 GCTACCACTCTTCCCAAGAGCAGAGCCGCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 3859

Qy 1801 ACTCGTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAAGGATATGCCTGA 1860  
Db 3860 ACTCGTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAAGGATATGCCTGA 3919  
Qy 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTTNAAGGNAAGCAGCAGAGAGAGCY 1920  
Db 3920 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTTNAAGGNAAGCAGCAGAGAGAGCY 3976  
Qy 1921 TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTNGGTTCNCCCGAGNAGACA 1980  
Db 3977 TCCGGCCCCCAGNCCCACTNATCAGTGTGT-----CCAGCGTGTGTTCNCCCGAGNAGACA 4032  
Qy 1981 GCTNACNATCANCACTGACATNCACTTCCCTTNGGCCCTTCCCTTNGGCCCANAGGATCTG 2040  
Db 4033 -----TCAGCATCACATGACATCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4083  
Qy 2041 CCGNACCGCACTTTGCACTCTGATGNAACCTCAAGCAGCTTTCATGGCTNGCCCTCTTNG 2100  
Db 4084 CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGCAGCTTTCATGGCTGCTGCTGCTGCTG 4140  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAAGCCAGAGATGG 2160  
Db 4141 GGGCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAAGCCAGAGATGG 4190  
Qy 2161 GGTNAAAGGNCAGTCTTGGAGCTGTCCANCACTGATGATGCTTCAAACTCTTNN 2220  
Db 4191 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4245  
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCTTNGGCCCTTNGGCCCTTNGGCCCTTNGGCCCTTNG 2280  
Db 4246 --CCAGATTTCTTAAGATAGCANCCTTNGGCCCTTNGGCCCTTNGGCCCTTNGGCCCTTNG 4301  
Qy 2281 CCAGGGGAGCTNCTCAGGACTCAGTGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340  
Db 4302 CCAGGGGAGCTA-CTCAGGACTCAGTGCATTAATCAGCTGTG-AATCGTCAGGGGG 4359  
Qy 2341 TGCTGTAGCTCAACCTCTCGGGCAGGGGAGCGCCAGACTCCGTGGGAGAGCTCAT 2400  
Db 4360 TGCTGTAGCTCAACCTCTCGGGCAGGGGAGCGCCAGACTCCGTGGGAGAGCTCAT 4419  
Qy 2401 TCCACATCTTGGCAGAGAGCTTTNGTCCAGCTGTCCAGTGTCCATGAGTCAAGCTGCTCC 2460  
Db 4420 TCCACATCTTGGCAGAGAGCTTT--GTCCAGCTGTCCATGAGTCAAGCTGCTCC 4478  
Qy 2461 GGGGAGAGAGCCCGCCCCCAGCAGATTAAGACTGCAGCTGTGCTGCTGCTGCTGCTGCTG 2520  
Db 4479 GGGGAGAGAGCCCGCCCCCAGCAGATTAAGACTGCAGCTGTGCTGCTGCTGCTGCTGCTG 4538  
Qy 2521 GGTGTGAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTAAAAA 2578  
Db 4539 GGTGTGAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTAAAAA 4596

## RESULT 4

AX774978 LOCUS AX774978 4596 bp DNA linear PAT 09-JUL-2003  
DEFINITION Sequence 294 from Patent WO03038129.  
ACCESSION AX774978  
VERSION AX774978.1 GI:32486494  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
MAMMALIA; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
Raponi, M.  
AUTHORS Methods for assessing and treating leukemia  
TITLE Patent: WO 03038129-A 294 08-MAY-2003;  
JOURNAL Ortho-Clinical Diagnostics, Inc. (US)  
FEATURES  
Location/Qualifiers  
source 1. 4596  
/organism="Homo sapiens"

/mol_type="unassigned DNA" /db_xref="taxon:9606"		Db		3060		TGAGAAACAGGCGCTAACCTGCGCTCCACCGCGCTCCACACTGCCGAAAGCAGCCT		3119			
Query Match		Qy		1021		TCCTGCTCGGTGCA CGATGCTGCCCTGTA AAAACACAGGCTCAGCGCTTCCCGGGATYTG		1080			
Best Local Similarity		Db		3120		TCCTGCTCGGTGCA CGATGCTGCCCTGTA AAAACACAGGCTCAGCGCTTCCCGGGATYTG		3177			
Matches 2334; Conservative		Qy		1081		NCCAGCCCCCGGCTCARGNCTGGAACACAGGCGCTCGNACAGNAGNAAAGTNGGGG		1140			
		Db		3178		GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGGCTCG-----CAGCAGCAAGGTGGG		3230			
		Qy		1141		CAAGCAGAGATGCTCCAGGATTTTCAACNCTGAGCCCNVTCGCCCANCCCTGCTGAADA		1200			
		Db		3231		GCAAGCAGAGATGCTCCAGGATTTTCAACCTGAGCCCTGCCAC-----CCTGCTGA		3284			
		Qy		1201		AAACAYTNCGCCACAGTGAAGACAGAGGAGGATGNCAGGATTTNACCTYGGGGAA		1260			
		Db		3285		AAAAACATCCGCCACAGTGAAGACAGAGGAGGATGNC-----AGGATTTACCTGGGGAA		3340			
		Qy		1261		ACAAACAGGAGTCTTTTCTGCTCCCTGCTCCAGTNCAGATTTGGCTGNAACCCGTTGG		1320			
		Db		3341		ACAAACAGGAGTCTTTTCTGCTCCCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG		3396			
		Qy		1321		ANTCAGTGACATTTGTTGGCAGACAGGGGAGAGAGCTTCCAGCTTGGGTGAGAGGG		1380			
		Db		3397		GATCAGTGACATTTGTTGGCAGACAGGGGAGAGAGCTTCCAGCTTGGGTGAGAGGG		3455			
		Qy		1381		GTGGGCGAGCCCTCGGCCCTTCAACCTTCCAGTNCAGATTTGGCTGNAACCTGTTGTA		1440			
		Db		3456		GTGGGCGAGCCCTCGGCCCTTCAACCT--CAAGCTGCTG--AGAGTGTCAAGTGTGTA		3513			
		Qy		1441		AGGNCACAAANTCAGANTTCAAGTGCAGAAACAGGTCAGAGGATTTGCCCGCCGNTA		1500			
		Db		3514		AGGCCCCAAATC---AGGTTCAAGTGCAGAAACAGGTCAGAGGATTTGCCCGCCGNTA		3569			
		Qy		1501		GTTTANNNGGGGCGCTCTTAAACCCCTTGCTGGCTGCTTCAACCTTGGCTGCTTCAACCC		1560			
		Db		3570		GTTAA---GGGGGCGCTCTTAAACCCCTTGCTGGCTGCTTCAACCTTGGCTGCTTCAACCC		3621			
		Qy		1561		CTTTTGGGTGAGGGGAAAGAAATGCTTCAACCTTGGGAAAGCTTCCCTGCTGAGAAATAC		1620			
		Db		3622		CTTTTGGGTGAGGGGAAAGAAATGCTTCAACCTTGGGAAAGCTTCCCTGCTGAGAAATAC		3680			
		Qy		1621		CACACTTTTCAGGTTGTCACACAGTCTGAGTTGACCTCTGTTTCAAGGAGGAC		1680			
		Db		3681		CACACTTTTCAGGTTGTCACACAGTCTGAGTTGACCTCTGTTTCAAGGAGGAC		3740			
		Qy		1681		AAAGAGGTGTAAAGTGAAGTGTCTCAGTNCCTCAGACATGTGCCCCCTTTGCTGTG		1740			
		Db		3741		AAAGAGGTGTAAAGTGAAGTGTCTCAGTNCCTCAGACATGTGCCCCCTTTGCTGTG		3799			
		Qy		1741		GCTACACTCTTCCCGAGCAGCAGGCCCCGAGCCCCCTTCAAGGCCAGCATGCCCCCAG		1800			
		Db		3800		GCTACACTCTTCCCGAGCAGCAGGCCCCGAGCCCCCTTCAAGGCCAGCATGCCCCCAG		3859			
		Qy		1801		ACTCGTGCACTCAGTTCCCTCATCTTAAGGTGAAGGTGATGCAAGGATATGCTTGA		1860			
		Db		3860		ACTCGTGCACTCAGTTCCCTCATCTTAAGGTGAAGGTGATGCAAGGATATGCTTGA		3919			
		Qy		1861		CAGGAACAGTCTGTGATGGAATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGAGAG		1920			
		Db		3920		CAGGAACAGTCTGTGATGGAATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGAGAG		3976			
		Qy		1921		TCCGGCGCCCCAGNCCCCACTNATCAGTGTNCCAGGCTGCTGTTNCCCGAGNAGACA		1980			
		Db		3977		TCCGGCGCCCCAGNCCCCACTNATCAGTGTNCCAGGCTGCTGTTNCCCGAGNAGACA		4032			
		Qy		1981		GCTNAGNATCANCACTGACACTNACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG		2040			
		Db		4033		-----TCAAGCATCACTGACACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG		4083			
		Qy		2041		CCGACGCGCACTTTGACATCTGATGATCCTCAAGACACTTTCATGCTGCTGCTGCTGCTG		2100			



[illegible][illegible]



QY 541 TCTCCCTCTCGGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAGCTCGCGGACA 600  
 Db 2478 TCTCCCTCTCGGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAGCTCGCGGACA 2537  
 QY 601 CCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTTCCAGCTTGA 660  
 Db 2538 CCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTTCCAGCTTGA 2597  
 QY 661 ACATGGTGTGGCCCGGGGGGGCCACCCGACACCCCAAGCTATTTCATAGGTGTGAAG 720  
 Db 2598 ACATGGTGTGGCCCGGGGGGGCCACCCGACACCCCAAGCTATTTCATAGGTGTGAAG 2657  
 QY 721 TCCAAATACAGTCTCTTAATGGTGAACACTTGCATCCCGGAGTTCACCGGGTCAAG 780  
 Db 2658 TCCAAATACAGTCTCTTAATGGTGAACACTTGCATCCCGGAGTTCACCGGGTCAAG 2717  
 QY 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTGGTCA 840  
 Db 2718 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTGGTGA 2777  
 QY 841 CCAAGAGCGGCGAGCTGTTCGTACGATGGAGGTGCCAGCTCGGGCATTCGACTGCG 900  
 Db 2778 CCAAGAGCGGCGAGCTGTTCGTACGATGGAGGTGCCAGCTTCGAGCTGCGACTGCG 2837  
 QY 901 AGTGACACTGCGCCCTGATGACAGCTTCGCTGAGCTGGAGGTCAAGCATGSCCAGC 960  
 Db 2838 AGTGACACTGCGCCCTGATGACAGCTTCGCTGAGCTGGAGGTCAAGCATGSCCAGC 2897  
 QY 961 TGAGAAACAGGCGCTTAACCTGCTCCACCCGCGCTCCACACTGCGCGGAAAGCAGCT 1020  
 Db 2898 TGAGAAACAGGCGCTTAACCTGCTCCACCCGCGCTCCACACTGCGCGG--AAGCAGCT 2956  
 QY 1021 TCTGTCTGGTGACGATGCTGCTGAAGAAACACAGGCTCAGCCGTTCCAGGGATATG 1080  
 Db 2957 TCTGTCTGGTGACGATGCTGCTG--AAACACAGGCTCAGCCGTTCCAGGG--ATC 3012  
 QY 1081 NCCAGCCCCCGCTCAGCAGTGGAAACAGGCGCTCGNAGCAGNAGCAGTGGGGG 1140  
 Db 3013 TGCCAGCCCCCGCTCAGCAGTGGGACAGGGGCTCGAGCAGCAAG-----GTGGGG 3065  
 QY 1141 CAAAGNAGATGCTCCCGAGATTTCCANCTGAGCCNTGCCANCCCTGCTGAADA 1200  
 Db 3066 GCAAGCAGATGCTCCCGAGATTTCCACC--TGAGCCCTGCCCCACCCCTGCTGAGAA 3122  
 QY 1201 AACAATNCCGACGCTGAAGACAGAGAGGATGAGCAGGATGAGCTTNNACCTTGGGAA 1260  
 Db 3123 AACACT--CGGCCACGTGAAGACACA---GAGGAGGATGGCAGGATTAACCTCGGAAA 3176  
 QY 1261 ACAAACAGGATCTTTNTTCTGCCCTGCTCCAGTNCAGTGGCTGNAACCCGCTTGG 1320  
 Db 3177 CAAACAGGATCTTCTGCTGCCCTGCTCCAG-----TCGAGTTGGCTGACCCGCTT--G 3227  
 QY 1321 ANTCAGTGACCATTTGTGGCAGANCGGGGAGAGAGCTTCCAGCTCGGTGAGAGGG 1380  
 Db 3228 GATCAGTGAACATTTGTGGCAGA-CAGGGGAGAGAGCTTCCAGCTCGGTGAGAGGG 3286  
 QY 1381 GTGGCGAGCCCTTCGGCCCTCACCCTNCCAGCTGCTGTGNAGAGTGTCAAGTGTGA 1440  
 Db 3287 GTGGCGAGCCCTTCGGCCCTCACCCT--CCAGCTGCTGTG-AGAGTGTCAAGTGTGA 3344  
 QY 1441 AGGNCNCCAAANTCAGGNTTCAGTGCAGAACAGGTCNAGCAGGATATGCCGCCCGNTA 1500  
 Db 3345 AGGCCCAAACTC--AGGTTCAAGTGAGAACACAGGT-CAGCAGGATATGCCGCCCGTAG 3400  
 QY 1501 GGTAAANNNGGGGCTCTWAACCCCTTTCCTTNGGCTNCACTNGGCCAGCTCANCC 1560  
 Db 3401 GTTAA----GGGGGCTCTTAAACCCCTTTCCTTNGGCTCA--CCTGGCCAGCTCA-CCC 3452  
 QY 1561 CTTTGTGGGTGAGGGGAAAGATGCTGACCTTGGGAAGGCTWCCCTGGTAGAATACAC 1620  
 Db 3453 CTTTGTGGGTGAGGGGAAAGATGCTGACCTTGGGAAGGCT-CCCTGGTAGAATACAC 3511  
 QY 1621 CACACTTTTCAAGTTGTGMAACACAGGTCTCGATTGACCTCTGCTTACGCCAAGGACC 1680

Db 3512 CACACTTTTCAAGTTGTGMAACACAGGTCTCTGCTGAGTTGACCTCTGCTTTCAGCCAAAGGACC 3571  
 QY 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCACAGACATGTGCCCTTTGCTGCTG 1740  
 Db 3572 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTGCTGCTG 3630  
 QY 1741 GCTTACCACTTTTCCACAGAGCAGCGGCCCGAGGCCCTTCAGGGCCAGACACTGCCCCAG 1800  
 Db 3631 GCTTACCACTTTTCCACAGAGCAGCGGCCCGAGGCCCTTCAGGGCCAGACACTGCCCCAG 3690  
 QY 1801 ACTCGCTGGCAGCTCAGTTCCTCATCTGTAAAGGTGAAGGTGATGCGAGATATGCTGA 1860  
 Db 3691 ACTCGCTGGCAGCTCAGTTCCTCATCTGTAAAGGTGAAGGTGATGCGAGATATGCTGA 3750  
 QY 1861 CAGGAACAGTCTGTGATGGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1920  
 Db 3751 CAGGAACAGTCTGTGATGGACATGATGATGATGATGATGATGATGATGATGATGATGATG 3808  
 QY 1921 TCCGGGGCCCCCAGNCCCACTNATCAGTGTNCCAGGCTGCTGCTGCTGCTGCTGCTGCTG 1980  
 Db 3809 TCCGGGGCCCCCAGNCCCACTNATCAGTGT--CCAGGCTGCTGCTGCTGCTGCTGCTG 3857  
 QY 1981 GCTNCAAGNCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2040  
 Db 3858 GCACAGCTCAGCATCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3915  
 QY 2041 CCGNACGGCACTTTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100  
 Db 3916 CCG-ACGGCACTTTGCACT--CTGATGATGATGATGATGATGATGATGATGATGATGATG 3972  
 QY 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTGTAGNAGCATANGCANGCANGCANGCANGC 2160  
 Db 3973 GGGCAGGCTAGGGCAGTGCACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4022  
 QY 2161 GGTGNAAGGNCAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
 Db 4023 GGTGAA--GGGACACAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 4077  
 QY 2221 NCCAGNATTTCTTAAGATAGCANCCCCCTTNGCCCATTTGCCCCAGCTTTAGCCTTTCT 2280  
 Db 4078 --CCAGATTTCTTAAGATAGCACCCCC--TTCCCCATTTGCCCCAGCTTTAGCCTTTCT 4133  
 QY 2281 CCCAGGGAGTANCTCAGGACTCAGTAGCATTAATTAATTAATTAATTAATTAATTAATTA 2340  
 Db 4134 CCCAGGGAGCTA-CTCAGGACTCAGTAGCATTAATTAATTAATTAATTAATTAATTAATTA 4191  
 QY 2341 TGTCTGTAGCTCAACTCTCTGGGCGAGGCGCGAGACTCCGTGGGAGAGAGCTCAT 2400  
 Db 4192 TGTCTGTAGCTCAACTCTCTGGGCGAGGCGCGAGACTCCGTGGGAGAGAGCTCAT 4251  
 QY 2401 TCCACATCTTTGCCAAGACAGCTTTNGTCCAGTGTCCCATTTGATGATGATGATGATGATG 2460  
 Db 4252 TCCACATCTTTGCCAAGACAGCTTT--GTCCAGCTGTCCCATTTGATGATGATGATGATG 4310  
 QY 2461 GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGAGCTTGGTACTGAGAGTCTG 2520  
 Db 4311 GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGAGCTTGGTACTGAGAGTCTG 4370  
 QY 2521 GGTCTGTAGAGACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAAAA 2579  
 Db 4371 GGTCTGTAGAGACTCTTTGTAAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAAAA 4429

RESULT 7  
 CO723104  
 LOCUS  
 DEFINITION Sequence 9038 from Patent WO02068579.  
 ACCESSION CO723104  
 VERSION CO723104.1 GI:42283961  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
AUTHORS  
Kits, such as nucleic acid arrays, comprising a majority of  
TITLE  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL  
Patent: WO 02068579-A 9038 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Location/Qualifiers  
source  
1..4584  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 67.7%; Score 1781.8; DB 6; Length 4584;  
Best Local Similarity 89.3%; Pred. No. 0;  
Matches 2296; Conservative 5; Mismatches 184; Indels 87; Gaps 35;  
Qy 1 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60  
Db 2100 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159  
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
Db 2160 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2219  
Qy 121 GCCCTTGGAGGGGAGATATAAGAAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 180  
Db 2220 GCCCTTGGAGGGGAGATATAAGAAACCAAGACATCCACGCCAAATCAAGCCAAATTACC 2279  
Qy 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCGAGGGCCCGGCCAG 240  
Db 2280 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCGAGGGCCCGGCCAG 2339  
Qy 241 CTGAGGAGACAAGGAGAGCCCTTAAGTCTCAGCTCTCTCCCAACAGAGCCCGCCAG 300  
Db 2340 CTGAGGAGACAAGGAGAGCCCTTAAGTCTCAGCTCTCTCCCAACAGAGCCCGCCAG 2399  
Qy 301 AGCCAAACAAGTCTCCCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCTT 360  
Db 2400 AGCCAAACAAGTCTCCCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCTT 2459  
Qy 361 TACCTCTGTCTCCTCGAGCCAGCCCTCGCAGAAACCCAGCTCACCAGAGCGGAAAG 420  
Db 2460 TACCTCTGTCTCCTCGAGCCAGCCCTCGCAGAAACCCAGCTCACCAGAGCGGAAAG 2519  
Qy 421 CAACCGTCCCGAGCAGGAACTGACAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCCTGT 480  
Db 2520 CAACCGTCCCGAGCAGGAACTGACAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCCTGT 2579  
Qy 481 CCAGGCCATTTCTCTGAGGAGCAGAGCAAAATTCCTGTGCTCAGCATCGACAGCC 540  
Db 2580 CCAGGCCATTTCTCTGAGGAGCAGGAGCAAAATTCCTGTGCTCAGCATCGACAGCC 2639  
Qy 541 TCTCCCTGTGGATGACAGTGAGAGAACCCATCAAGAGGCTCTCAAGCTTCGGGGACA 600  
Db 2640 TCTCCCTGTGGATGACAGTGAGAGAACCCATCAAGAGGCTCTCAAGCTTCGGGGACA 2699  
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 660  
Db 2700 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 2759  
Qy 661 ACATGGTGTGGCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720  
Db 2760 ACATGGTGTGGCCGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2819  
Qy 721 TCCAAATACAGTCTCTTAAATGGTGAAACACCTGACATCCGGGAGTTCACCGGGTCAAG 780  
Db 2820 TCCAAATACAGTCTCTTAAATGGTGAAACACCTGACATCCGGGAGTTCACCGGGTCAAG 2879  
Qy 781 TGGAGACATCGGCACCTGGCATCAGCAGCCAGATCCCGAGCTGCAGCCTTCAGCTTGGTCA 840

Db 2880 TGGAGACATCGGCACCTGGCATCAGACCCAGATCCCGAGCTTCAGCTTGGTCA 2939  
Qy 841 CCAAAGACGGGACGCTGTTTCGCTACACATGAGAGTGCAGAGTTCGGGCAATCAGCTGC 900  
Db 2940 CCAAAGACGGGACGCTGTTTCGCTACACATGAGAGTGCAGAGTTCGGGCAATCAGCTGC 2999  
Qy 901 AGTGACACACTGGCCCTTGATGGCAGCTTCGCTTGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTGACACACTGGCCCTTGATGGCAGCTTCGCTTGAGCTGGAGGTCAAGCATGGCCAGC 3059  
Qy 961 TGGAGAACAGGCCCTTAACCTTCGCCCTCACCGCCGCTCCACACTGCGGGAAGCAGCCT 1020  
Db 3060 TGGAGAACAGGCCCTTAACCTTCGCCCTCACCGCCGCTCCACACTGCGG--AAGCAGCCT 3118  
Qy 1021 TCCTGCTCGGTGACACGATGTCGCCCTGAAAACACAGCTCAGCCGTTTCCCAGGGGATVTG 1080  
Db 3119 TCCTGCTCGGTGACACGATGTCGCCCTG--AAACACAGGCTCAGCCGTTTCCCAGGG--ATC 3174  
Qy 1081 NCCAGCCCCCGGCTCARGAGTGGGAAACAGGCGCTTCGNCAGCAGNAGNAGTNGGGGG 1140  
Db 3175 TGCAGCCCCCGGCTCAGCAGTGGGACCCAGGGCTTCGACAGCAGCAAG-----GTGGGG 3227  
Qy 1141 CAAGCNAAGATGCTCTCCAGGATTTTCAANCTGAGCCCTNTGCCANCCCTGCTGAADA 1200  
Db 3228 GCAAGCAGAAATGCTCTCCAGGATTTTCAAC--TGAGCCCTGCCCCACCTGCTGTGAAA 3284  
Qy 1201 AAACAYTNCGCCACGCTGAAGACAGAGAGGATGNCAGGAGTTNNACCTTGGGGAA 1260  
Db 3285 AACACT--CCGCCACGTGAAGAGACA---GAGGAGATGGCAGGAGTTACTCGGGAAA 3338  
Qy 1261 ACAAACAGGGATCTTTNTTCTGCCCTGTCTCAGTNCGAGTTGGCTTCGACCCGCTTGG 1320  
Db 3339 CAACAGGATCTTCTGCCCCCTGCTCCAG-----TCGAGTTGCGCTGACCCGCTTGG- 3390  
Qy 1321 ANTCACTGACATTTGTTTGGCAGANCAAGGGGAGAGAGCTTTCAGCTTGGGTCAAGAGGG 1380  
Db 3391 -ATCAGTGAACATTTGCTGGCAGA-CAGGGGAGAGAGCTTCCAGCTTGGGTCAAGAGGG 3448  
Qy 1381 GTGGCGAGCCCTTCGGCCCTCACCTTNCAGGCTCTGTGNAGAGTGTCAAGTGTGA 1440  
Db 3449 GTGGCGAGCCCTTCGGCCCTCACCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 3506  
Qy 1441 AGGNCNCAANCTCAGNTTCAGTGCAGAACACAGGTNCAAGCAGTATGCCGCCCGNTA 1500  
Db 3507 AGGCCCAAACTC---AGGTTCACTGCAAGAACAGGT-CAGCAGTATGCCGCCCGCTAG 3562  
Qy 1501 GGTAAANNNGGGGCGCTCTNAAACCCCTTCCTTNGGCTTCACTTNGGCCAGCTCAACCC 1560  
Db 3563 GTTAA-----GGGGCCCTCTAAACCCCTTCCTGGCCCTCA---CCTGGCCAGCTCA-CCC 3614  
Qy 1561 CTTTGGGTGTAGGGGAAAGAAATGCTGACCTTGGGAAGGCTTCCCTGGTAGAATACAC 1620  
Db 3615 CTTTGGGTGTAGGGGAAAGAAATGCTGACCTTGGGAAGGCT-CCCTGGTAGAATACAC 3673  
Qy 1621 CACACTTTTTCAGGTGTGTCACACAGGTCTCTGAGTTGACTCTGTGTTTTCAGCCAGGACC 1680  
Db 3674 CACACTTTTTCAGGTGTGTCACACAGGTCTCTGAGTTGACTCTGTGTTTTCAGCCAGGACC 3733  
Qy 1681 AAAGAAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGTGCCCTTTGCTGCTG 1740  
Db 3734 AAAGAAAGGTGTGAAGTGAAGTGTCTCAGT-CCCCAGACATGTGTGCCCTTTGCTGCTG 3792  
Qy 1741 GCTACACATCTTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCAGCAGACTGCCCCAG 1800  
Db 3793 GCTACACATCTTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCAGCAGACTGCCCCAG 3852  
Qy 1801 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGAGGTGATGAGATATGCTGTA 1860  
Db 3853 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGAGGTGATGAGATATGCTGTA 3912  
Qy 1861 CAGGAACAGCTCTCTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGACGY 1920







/clone_lib="TESB2"	
/note="cloning vector: pME18SFL3"	
218.1651	
/note="unnamed protein product"	
/codon_start=1	
/translation="MKLIRSSSCHVGSARKLHWRAGOTEPGWAGTRHDSVPVLPPIRG	
/db_xref="GI:47077392"	
/protein_id="BAD18584.1"	
GLADNVLLSDGSHALCDRGHVCLOPDGLGKSLITGDYIPGTEPHMAPEVVLGRS	
CDKVDVWSSCCMLHMLNGCHPWTQFFRGPLCLKIASEPFPVPRPIPPSCAPLTAQAI	
OEGURKEPIHVSAAELGKYNRALQVGGILKSPWRGEYKEPRHPHPNQANPLTHA	
QPRELSPAPRPAEETGRAPKLQPLPPEPEPNKSPPLTILSKESGNWEPILPS	
SLEPAPANPSPKKATVPEQLQLEIELFLNSLOPFSLEEQEQLISLISLISLS	
LSDSEKNPSKASQSSRDTLSSGVHSSQAEARSSNMWVLARGRPDTPSYFNGVK	
VQIQSLNGEHLHIREPHRVKVDIATGLISSQIPAAASLVTKDQPVRYDNEVPDSGI	
DLQCTLAPDGSFAWSVRVKGQLENRP"	
ORIGIN	
Query Match	67.3%; Score 1771.6; DB 9; Length 3152;
Best Local Similarity	89.2%; Pred. No. 0;
Matches 2187; Conservative	5; Mismatches 186; Indels 87; Gaps 35;
QY	1 CCCCTCTCACAGCCAGGCGCATCCAGAGGGGCTCAGGAAAGAGCCCATCCACCGCGTGT 60
DB	
QY	675 CCCCTCTCACAGCCAGGCGCATCCAGAGGGGCTCAGGAAAGAGCCCATCCACCGCGTGT 734
DB	
QY	61 CTGCAGCGGAGTGGAGGAGGTGAACCGGCACTACGCAAGTGGAGGTCTGAAGA 120
DB	
QY	735 CTGCAGCGGAGTGGAGGAGGTGAACCGGCACTACGCAAGTGGAGGTCTGAAGA 794
DB	
QY	121 GCCCTTGGAGGGGAGATATAAGAAACCAAGACATCCACCGCCAATCAAGCCAATTACC 180
DB	
QY	795 GCCCTTGGAGGGGAGATATAAGAAACCAAGACATCCACCGCCAATCAAGCCAATTACC 854
DB	
QY	181 ACCAGACCTCTCATGCCAGCCGAGAGAGTTCGCGCAAGGGCCCGAGGCCCGGCCAG 240
DB	
QY	855 ACCAGACCTCTCATGCCAGCCGAGAGAGTTCGCGCAAGGGCCCGAGGCCCGGCCAG 914
DB	
QY	241 CTGAGGAGACACAGGAGCCCTTAGCTCCAGCCCTCTCTCCACAGAGCCCGCCAG 300
DB	
QY	915 CTGAGGAGACACAGGAGCCCTTAGCTCCAGCCCTCTCTCCACAGAGCCCGCCAG 974
DB	
QY	301 AGCCAAACAAGTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 360
DB	
QY	975 AGCCAAACAAGTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 1034
DB	
QY	361 TACCTCTGTCTCTCTGTGAGAGCAGCCCTTGCCAGAAACCCCAAGCTCACAGAGCGGAAG 420
DB	
QY	1035 TACCTCTGTCTCTCTGTGAGAGCAGCCCTTGCCAGAAACCCCAAGCTCACAGAGCGGAAG 1094
DB	
QY	421 CAACCGTCCGGAGCAGGAACTGACGAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 480
DB	
QY	1095 CAACCGTCCGGAGCAGGAACTGACGAGCTGGAATAGAAATATTCTCTCAACAGCCTGT 1154
DB	
QY	481 CCCAGCATTCTCTCTGAGGAGCAGGAGCAAAATCTCTGTCGCTCTCAGCATCGACGCC 540
DB	
QY	1155 CCCAGCATTCTCTCTGAGGAGCAGGAGCAAAATCTCTGTCGCTCTCAGCATCGACGCC 1214
DB	
QY	541 TCTCCCTGTGCGATGACAGTAGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 600
DB	
QY	1215 TCTCCCTGTGCGATGACAGTAGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA 1274
DB	
QY	601 CCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 660
DB	
QY	1275 CCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTGGA 1334
DB	
QY	661 ACATGTGTGTGGCCCGGGGGCGGCCCAACGACACCCCAAGCTATTTCATGTGTGAAAG 720
DB	
QY	1335 ACATGTGTGTGGCCCGGGGGCGGCCCAACGACACCCCAAGCTATTTCATGTGTGAAAG 1394
DB	
QY	721 TCCAAATACAGTCTCTTAATGTGGAACACCTGCACATCCGGAGGTTCCACCGGGTCAAAG 780
DB	
QY	1395 TCCAAATACAGTCTCTTAATGTGGAACACCTGCACATCCGGAGGTTCCACCGGGTCAAAG 1454
DB	
QY	781 TGGGAGACATCGCCACCTGGCATCAGCAGCCAGATCCAGCTTGCAGCTTGCAGTTGGTCA 840
DB	
QY	1455 TGGGAGACATCGCCACCTGGCATCAGCAGCCAGATCCAGCTTGCAGCTTGCAGTTGGTCA 1514
DB	
QY	841 CCAAAGACGGGACGCTGTTTGGCTTACACATGAGAGGTGCAGACTCGGGCATACACCTGC 900
DB	
QY	1515 CCAAAGACGGGACGCTGTTTGGCTTACACATGAGAGGTGCAGACTCGGGCATACACCTGC 1574
DB	
QY	901 AGTCACACTCTGGCCCTGATGAGGAGCTTCGCTCGAGCTGAGAGGTCAAGCATGGCCAGC 960
DB	
QY	1575 AGTCACACTCTGGCCCTGATGAGGAGCTTCGCTCGAGCTGAGAGGTCAAGCATGGCCAGC 1634
DB	
QY	961 TGGAGAACAGAGCCCTAAACCTTGCCTCCACCGCGGCTTCCACACTGCGCGG-AAAGAGCCT 1020
DB	
QY	1635 TGGAGAACAGAGCCCTAAACCTTGCCTCCACCGCGGCTTCCACACTGCGCGG-AAAGAGCCT 1693
DB	
QY	1021 TCCTGTCTGTGTCACAGTGTGCTCCTGAAACACAGAGCTCAGCCGTTCCTCCAGGGGATYTG 1080
DB	
QY	1694 TCCTGTCTGTGTCACAGTGTGCTCCTGAAACACAGAGCTCAGCCGTTCCTCCAGGG-ATC 1749
DB	
QY	1081 NCCAGCCCCCGGCTCARGAGTGGGAACAGGGCCCTCGNCAGCAGNAGCNAAGGTNGGGG 1140
DB	
QY	1750 TGCCAGCCCCCGGCTCAGCAGTGGGAACAGGGCCCTCGCAGCAGCAAG-----GTGGGG 1802
DB	
QY	1141 CAAGCNAAGATGCTCTCCAGGATTTTACANCTTGAGCCCTGAGCCCTGAGCCCTGCTGAADA 1200
DB	
QY	1803 GCAAGCAGATGCTCTCCAGGATTTTACACC---TGAGCCCTGCCCCCACCCTGCTGAGAA 1859
DB	
QY	1201 AAACAYNVCCGCACTGTAAGAGACAGAGAGAGATGNCAGAGATTTNNACTTGGGGAA 1260
DB	
QY	1860 AACACT--CCGCCACGTGAAGAGACA----GAGGAGGATGCGCAGAGTTACTCGGGAAA 1913
DB	
QY	1261 ACAAAACAGGGATCTTTNTTCTGCCCCCTGCTCCAGTNCAGTTCGCCCTGNAACCCGCTTGG 1320
DB	
QY	1914 CAACAGAGATCTTCTGCCCCCTGCTCCAG-----TCAGTGTGGCTTGAACCGCTT--G 1964
DB	
QY	1321 ANTCACTGACCATTTGTTGGCAGANCAGGGGAGAGCAGCTTCCAGCTTGGGTTCAGAAAGG 1380
DB	
QY	1965 GATCAGTGACCATTTGCTGGCAGA-CAGGGGAGAGCAGCTTCCAGCTTGGGTTCAGAAAGG 2023
DB	
QY	1381 GTGGGCGAGCCTTCGGCCCTTCACTCCCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440
DB	
QY	2024 GTGGGCGAGCCTTCGGCCCTTCACTCCCT--CCAGGCTGCTGTG-AGAGTGTCAAGTGTGTA 2081
DB	
QY	1441 AGGNCACAAANTCAGGNTTTCAGTGCAGAACACAGTNCAGCAGGTATGCCGCCGNTA 1500
DB	
QY	2082 AGGGCCCAATC---AGTTTCAGTGCAGAACACAGGT--CAGCAGGTATGCCGCCGCTAG 2137
DB	
QY	1501 GGTAAANNNGGGGGCCCTCTNAAACCCCTTGGCTNGGCCCTNCACTNGGCCAGCTCANCCC 1560
DB	
QY	2138 GTTAA---GGGGGCCCTTAAACCCCTTGGCTTGGCCCTCA---CTGGCCAGCTCA-CGC 2189
DB	
QY	1561 CTTTGGGTGTAGGGGAAAAGATGCTGACCTCGGGAAGGCTWCCCTGGTAGAATACAC 1620
DB	
QY	2190 CTTTGGGTGTAGGGGAAAAGATGCTGACCTCGGGAAGGCT--CCCTGGGTAGAAATACAC 2248
DB	
QY	1621 CACACTTTTTCAGTGTGTTGCAACAGAGTCTGAGTTCACCTCTCTGTTTCAGCCTCAAGGACC 1680
DB	
QY	2249 CACACTTTTTCAGTGTGTTGCAACAGAGTCTGAGTTCACCTCTCTGTTTCAGCCTCAAGGACC 2308
DB	
QY	1681 AAAGAAGGTGTAAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCCCTTTTGTGTGTG 1740
DB	
QY	2309 AAAGAAGGTGTAAAGTGAAGTGTCTCAGT--CCCCAGACATGTGCCCCCTTTTGTGTGTG 2367
DB	
QY	1741 GCTACACTCTTCCCAGAGCAGCAGCCCGAGCCCTTTCAGGCCACAGCTTGCCTCCAG 1800
DB	
QY	2368 GCTACACTCTTCCCAGAGCAGCAGCCCGAGCCCTTTCAGGCCACAGCTTGCCTCCAG 2427
DB	
QY	1801 ACTCCTGCGCACTCAGTTCCTCTCATCTGTAAAGGTGAAGGGTGTATGCGAGGATATGCTGA 1860
DB	
QY	2428 ACTCCTGCGCACTCAGTTCCTCTCATCTGTAAAGGTGAAGGGTGTATGCGAGGATATGCTGA 2487
DB	



```

2946 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGGCTCG-----CAGCAACAAGGTGGG 2998
1141 CAAGCNAGAATGCTCCAGGATTTACANCTGAGCCNTGCCCCANCCCTGCTGAADA 1200
2999 GCAAGCAGAATGCTCCAGGATTTCAACCTGAGCCCTGCCCCAC-----CCTGCTGA 3052
1201 AAACAATNCCGCGCAGCTGAAGAGACAGAAAGAGGATGNCAGGAGTTNNACTCTGGGAA 1260
3053 AAAAACAATCCGCGCAGCTGAAGAGACAGAAAGAGGATGSC---AGGAGTTACTGGGAA 3108
1261 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTTGGCCTG 1309
3109 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGT-CGAGTTGGCCTG 3155

RESULT 11
AR063065
LOCUS AR063065 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5844073.
ACCESSION AR063065
VERSION AR063065.1 GI:5990756
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 3156)
AUTHORS Rothe,M. and Wu,L.
TITLE Human NIK proteins
JOURNAL Patent: US 5844073-A 1 01-DEC-1998;
FEATURES
  Location/Qualifiers
    source
      1..3156
      /organism="unknown"
      /mol_type="unassigned DNA"

ORIGIN
Query Match 43.1%; Score 1133.4; DB 6; Length 3156;
Best Local Similarity 94.8%; Pred. No. 1.2e-279;
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;

QY 1 CCCCTCTACAGCCAGCCATCCAAAGAGGGGCTGAGGAAGAGGCCCATCCACCGCGTGT 60
DB 1868 CCCCTCTACAGCCAGCCATCCAAAGAGGGGCTGAGGAAGAGGCCCATCCACCGCGTGT 1927
QY 61 CTGACGCGAGCTGGAGGGAAGGTGAACCGGCGACTACAGCAAGTGGAGGTCTGAAGA 120
DB 1928 CTGACGCGAGCTGGAGGGAAGGTGAACCGGCGACTACAGCAAGTGGAGGTCTGAAGA 1987
QY 121 GCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180
DB 1988 GCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 2047
QY 181 ACCAGACCTTCATGCCAGCGAGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCCAG 240
DB 2048 ACCAGACCTTCATGCCAGCGAGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCCAG 2107
QY 241 CTGAGAGACAACAGCAGAGCCCTTAAGCTCAGCCTCTCTCCACAGAGCCCGCCAG 300
DB 2108 CTGAGAGACAACAGCAGAGCCCTTAAGCTCAGCCTCTCTCCACAGAGCCCGCCAG 2167
QY 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGAAGGAGAGTCTGGGATGTTGGGAACCCCT 360
DB 2168 AGCCAAACAAGTCTCTCCCTTGACTTTGAGAAGGAGAGTCTGGGATGTTGGGAACCCCT 2227
QY 361 TACCTCTGTCTCTCTGGAGCAGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAG 420
DB 2228 TACCTCTGTCTCTCTGGAGCAGCCCTGCGCAGAAACCCAGCTCACCAGAGCGGAAG 2287
QY 421 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATTCTTCAACAGCCTGT 480
DB 2288 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATTCTTCAACAGCCTGT 2347
QY 481 CCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTCTCTCTGCTCAGCATCGACGCC 540
```

```

2348 CCCAGCCATTTCTCTGGAGGAGCAGGAGAAATTTCTCTGCTGCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAGCTCGCGGGACA 600
DB 2408 TCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAAAGGCTCTCAAGCTCGCGGGACA 2467
QY 601 CCCTGAGCTCAGGCGGTACACTCTCTGGAGCAGCCAGGCGCAGAGCTCGAAGCTCCAGCTGGA 660
DB 2468 CCCTGAGCTCAGGCGGTACACTCTCTGGAGCAGCCAGGCGCAGAGCTCGAAGCTCCAGCTGGA 2527
QY 661 ACATGCTGCTGGCCCGGGGGCGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAG 720
DB 2528 ACATGCTGCTGGCCCGGGGGCGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGCTCAAG 780
DB 2588 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGCTCAAG 2647
QY 781 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGAGCTGCAGCCTTCAGCTTGTCA 840
DB 2648 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGAGCTGCAGCCTTCAGCTTGTCA 2707
QY 841 CCAAAGACGGGCGACGCTGTTGCTTACGACATGGAAGTGCAGACTCGGGCATCGACCTGC 900
DB 2708 CCAAAGACGGGCGACGCTGTTGCTTACGACATGGAAGTGCAGACTCGGGCATCGACCTGC 2767
QY 901 AGTGCACACTGGCCCTGATGGCAGCTTCGCTCGAGCTGAGGCTCAAGCATGGCCAGC 960
DB 2768 AGTGCACACTGGCCCTGATGGCAGCTTCGCTCGAGCTGAGGCTCAAGCATGGCCAGC 2827
QY 961 TGGGAAACAGGCGCCTAACCTTCGCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 1020
DB 2828 TGGGAAACAGGCGCCTAACCTTCGCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 2887
QY 1021 TCCTGCTCGTGCAGATGCTGCCCTGAAACAACAGGCTCAGCGCTTCCAGGGGATYTG 1080
DB 2888 TCCTGCTCGTGCAGATGCTGCCCTGAAACAACAGGCTCAGCGCTTCCAGGGGATYTG 2945
QY 1081 NCCAGCCCCCGGCTCAGCAGTGGGAACAGGCGCTCGNCAGCAGNAGNAAAGGTNGGGG 1140
DB 2946 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGGCTCG-----CCTGCTGA 2998
QY 1141 CAAGCAGAAATGCTCTCCAGGATTTCAACNCTGAGCCNTGCCCANCCCTGCTGAADA 1200
DB 2999 GCAAGCAGAAATGCTCTCCAGGATTTCAACNCTGAGCCCTGCCCCAC-----CCTGCTGA 3052
QY 1201 AAACAATNCCGCGCAGCTGAAGAGACAGAGAGGATGNCAGGAGTTNNACTCTGGGAA 1260
DB 3053 AAAAACAATCCGCGCAGCTGAAGAGACAGAAAGAGGATGSC---AGGAGTTACTGGGAA 3108
QY 1261 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTTGGCCTG 1309
DB 3109 ACAAAACAGGGATCTTTTCTGCCCCCTGCTCCAGT-CGAGTTGGCCTG 3155

RESULT 12
AR068515
LOCUS AR068515 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5854003.
ACCESSION AR068515
VERSION AR068515.1 GI:6000722
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 3156)
AUTHORS Rothe,M. and Wu,L.
TITLE Screening method for agents that modulate human NIK activity
JOURNAL Patent: US 5854003-A 1 29-DEC-1998;
FEATURES
  Location/Qualifiers
    source
      1..3156
      /organism="unknown"
```



QY	181	ACGAGACCTCCATGCCCCAGCGAGAGAGCTTTTCGCCAAGGGCCCCCGGCGAG	240
Db	2048	ACGAGACCTCCATGCCAGCGAGAGAGCTTTTCGCCAAGGGCCCCCGGCGAG	2107
QY	241	CTGAGGACACAGGAGAGAGCCCTTACGCTCCAGCCTCTCCACACAGAGCCCGAG	300
Db	2108	CTGAGGACACAGGAGAGAGCCCTTACGCTCCAGCCTCTCTCTCCACAGAGCCCGAG	2167
QY	301	AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGGGAACCT	360
Db	2168	AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGGGAACCT	2227
QY	361	TACCTCTGTCTCTCTCTGAGCAGAGCCCTTCCAGAGAACCCAGCTCCAGAGCGGAAG	420
Db	2228	TACCTCTGTCTCTCTCTGAGCAGAGCCCTTCCAGAGAACCCAGCTCCAGAGCGGAAG	2287
QY	421	CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTAATCTCAACAGCCTGT	480
Db	2288	CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTAATCTCAACAGCCTGT	2347
QY	481	CCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACGCC	540
Db	2348	CCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACGCC	2407
QY	541	TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA	600
Db	2408	TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA	2467
QY	601	CCTGAGCTCAGCGTACACTCTCTGGAGCAGCAGCGCCAGGCTCGAAGCTCCAGCTGGA	660
Db	2468	CCTGAGCTCAGCGTACACTCTCTGGAGCAGCAGCGCCAGGCTCGAAGCTCCAGCTGGA	2527
QY	661	ACATGTGTGTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCATGTGTGAAG	720
Db	2528	ACATGTGTGTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCATGTGTGAAG	2587
QY	721	TCCAAATACAGTCTCTTAATGTGACACTCGACATCCGGAGTTCCACCGGCTCAAAG	780
Db	2588	TCCAAATACAGTCTCTTAATGTGACACTCGACATCCGGAGTTCCACCGGCTCAAAG	2647
QY	781	TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA	840
Db	2648	TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA	2707
QY	841	CAAAAGCGGCGAGCTTTCGTACGATGAGAGTGCCAGACTCGGCGATCGACCTGC	900
Db	2708	CAAAAGCGGCGAGCTTTCGTACGATGAGAGTGCCAGACTCGGCGATCGACCTGC	2767
QY	901	AGTGCACTGGCCCTGATGGCAGCTTCGCTTGGAGCTGGAGGGTCAAGCATGGCCAGC	960
Db	2768	AGTGCACTGGCCCTGATGGCAGCTTCGCTTGGAGCTGGAGGGTCAAGCATGGCCAGC	2827
QY	961	TGGAGAACAGGCGCTTAACTGCTCCACCGCGGCTCCACACTCGCGGAAAGCAGCT	1020
Db	2828	TGGAGAACAGGCGCTTAACTGCTCCACCGCGGCTCCACACTCGCGGAAAGCAGCT	2887
QY	1021	TCTGTCTGCTGACGATGCTGCTTGAACACAGGCTCAGCGTTCCAGGGATYTG	1080
Db	2888	TCTGTCTGCTGACGATGCTGCTTGAACACAGGCTCAGCGTTCCAGGGATYTG	2945
QY	1081	NCCAGCCCCCGGCTCAGCAGNTGGGAAACAGAGGCTTCGNCAGCAGCAGGAGTNGGGG	1140
Db	2946	GCCAGCCCCCGGCTCA--CAGTGGGAAACAGAGGCTTCG--CAGCAGAGGTTGGG	2998
QY	1141	CAAGCAGAAATGCTCCAGGATTTCAACNCCTGAGCCCTGCCCCNCCCTCTGAADA	1200
Db	2999	GCAAGCAGAAATGCTCCAGGATTTCAACNCCTGAGCCCTGCCCCNCCCTCTGA	3052
QY	1201	AAACAYTNCGCCACGCTGAGAGACAGAGGATGNCAGGAGTTNNACCTYGGGAA	1260
Db	3053	AAACAYTNCGCCACGCTGAGAGACAGAGGATGNCAGGAGTTNNACCTYGGGAA	3108

QY	1261	ACAAAACAGGAGTCTTTNTTCTGCCCTGTCCAGTNCAGTTCGGCTG	1309
Db	3109	ACAAAACAGGAGTCTTT-TTCTGCCCTGTCTCCAGT-CGAGTTGGCTG	3155
RESULT 14			
LOCUS	AR429673	2844 bp	DNA linear PAT 18-DEC-2003
DEFINITION	Sequence 2 from patent US 6645728.		
ACCESSION	AR429673		
VERSION	AR429673.1	GI:40189976	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unkown.		
REFERENCE	1 (bases 1 to 2844)		
AUTHORS	Greene, W.C.; Lin, X. and Gelezuinas, R.		
TITLE	Inhibitor of the inflammatory response induced by TNF.alpha. and IL-1		
JOURNAL	Patent: US 6645728-A 2 11-NOV-2003;		
FEATURES	Location/Qualifiers		
source	1..2844		
ORIGIN	/organism="unkown"		
	/mol_type="genomic DNA"		
	Query Match 37.1%; Score 977; DB 6; Length 2844;		
	Best Local Similarity 100.0%; Pred. No. 1.8e-239;		
	Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	CCCCTCTCACAGCCAGGCCATCCAGAGGGCTGAGGAAAGAGCCCATCCACGCGTGT	60
Db	1868	CCCCTCTCACAGCCAGGCCATCCAGAGGGCTGAGGAAAGAGCCCATCCACGCGTGT	1927
QY	61	CTCAGCGGAGCTCGGAGGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA	120
Db	1928	CTCAGCGGAGCTCGGAGGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA	1987
QY	121	GCCTTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC	180
Db	1988	GCCTTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC	2047
QY	181	ACGAGACCTCCATGCCAGCGGAGAGCTTTCCGCAAGGGCCCGAGGCGCCGCGCAG	240
Db	2048	ACGAGACCTCCATGCCAGCGGAGAGCTTTCCGCAAGGGCCCGAGGCGCCGCGCAG	2107
QY	241	CTGAGGACACACAGGAGAGCGCCCTAAGCTCCAGCTTCTCTCCACAGAGCCCGCAG	300
Db	2108	CTGAGGACACACAGGAGAGCGCCCTAAGCTCCAGCTTCTCTCCACAGAGCCCGCAG	2167
QY	301	AGCCAAACAAGTCTCTCTCTGATTTGAGAGGAGAGTCTGGGATGGGACCCCT	360
Db	2168	AGCCAAACAAGTCTCTCTCTGATTTGAGAGGAGAGTCTGGGATGGGACCCCT	2227
QY	361	TACCTCTGTCTCTCTGAGCAGCGCCCTGCCAGAGAACCCAGCTCAGCAGCGGAAAG	420
Db	2228	TACCTCTGTCTCTCTGAGCAGCGCCCTGCCAGAGAACCCAGCTCAGCAGCGGAAAG	2287
QY	421	CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT	480
Db	2288	CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT	2347
QY	481	CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACGCC	540
Db	2348	CCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACGCC	2407
QY	541	TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA	600
Db	2408	TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACA	2467
QY	601	CCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCGCCAGGCTCGAAGCTCGAGCTGGA	660
Db	2468	CCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCGCCAGGCTCGAAGCTCGAGCTGGA	2527

```

Qy 661 ACATGGTCTGCGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCATCGTGTGAAG 720
Db 2528 ACATGGTCTGCGCCGGGGGGGGCCACCGACACCCCAAGCTATTTCATCGTGTGAAG 2587
Qy 721 TCCAAATACACTCTTAAATGGTGAACACCTGCACATCCGGGAGTTCACCGGTCAAAG 780
Db 2588 TCCAAATACACTCTTAAATGGTGAACACCTGCACATCCGGGAGTTCACCGGTCAAAG 2647
Qy 781 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTCA 840
Db 2648 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTCA 2707
Qy 841 CCAAAGACGGGAGCCTGTTTCGCTACGACATGGAGGTGCCAGATCGGGCATCGACCTGC 900
Db 2708 CCAAAGACGGGAGCCTGTTTCGCTACGACATGGAGGTGCCAGATCGGGCATCGACCTGC 2767
Qy 901 AGTCACACTGCGCCCTGATCGGAGCTTCGCTCGAGTGGAGGTCAAGCATGCCAGC 960
Db 2768 AGTCACACTGCGCCCTGATCGGAGCTTCGCTCGAGTGGAGGTCAAGCATGCCAGC 2827
Qy 961 TGGAGAACAGGCCCTAA 977
Db 2828 TGGAGAACAGGCCCTAA 2844

RESULT 15
AC003963 48032 bp DNA linear PRI 28-MAR-1998
LOCUS Homo sapiens chromosome 17, clone HRPCL169K15, complete sequence.
DEFINITION AC003963
ACCESSION AC003963.1 GI:2995604
VERSION HTG.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 48032)
AUTHORS Birren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P. and
Lander,E.
Homo sapiens chromosome 17, clone HRPCL169K15
Unpublished
REFERENCE 2 (bases 1 to 48032)
AUTHORS Birren,B., Fasman,K., McKernan,K., Munro,C., Nusbaum,C.,
Richardson,P., Lander,E., Allen,N., Baldwin,J., Barna,N.,
Beckerly,R., Cantu,C., Castle,A., Cooke,P., Daly,M.J., Depayre,E.,
Devon,K., Dewar,K., DuRette,B., Ferreira,P., Forrest,C., Gage,D.,
Gardyna,S., Gensheimer,S., Geraghty,K., Gilmartin,T., Gray,D.,
Hagos,B., Halphen,I., Harris,K., Horton,L., Howland,J.C., Huang,J.,
Hui,L., Jacotot,L., Linton,L., Mackenzie,J., Marquis,N.,
McDermott,J., McEwan,P., McGurk,A., Meldrim,J., Molla,M.,
Morris,W., Morrow,J., Nachman,A., Naylor,J., O'Connor,T.,
Pavlin,B., Peterson,K., Riley,R., Roberts,D., Rollins,G., Roy,A.,
Sarnaik,A., Shlu,P., Shyam,R., Stilwell,J., Stone,C.,
Strickland,C., Sydney,K., Tang,L., Vassiliev,H., Vo,A., Wu,Y.,
Ye,W.J., Zemtseva,I., Zhao,J. and Zody,M.
Direct Submission
Submitted (30-DEC-1997) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 48032)
AUTHORS Birren,B., Fasman,K., McKernan,K., Nusbaum,C., Richardson,P.,
Lander,E., Allen,N., Baker,J., Baldwin,J., Barna,N., Beckerly,R.,
Benn,J., Boatin,C., Boutwell,C., Brown,A., Byrne,S., Cantu,C.,
Castle,A., Cerny,J., Cooke,P., Daly,M.J., Depayre,E., Devon,K.,
Dewar,K., Donelan,L., DuRette,B., Etemadi,S., Ferreira,P.,
FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S.,
Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Gray,D.,
Hagos,B., Harris,K., Horton,L., Howland,J.C., Hui,L., Jacotot,L.,
Kann,L., Linton,L., Macdonald,P., Marquis,N., McEwan,P., McGurk,A.,
Meldrim,J., Molla,M., Morris,W., Morrow,J., Nachman,A., Naylor,J.,
O'Connor,T., Pavlin,B., Peterson,K., Riley,R., Roberts,D.,
Rollins,G., Rossello,R., Roy,A., Shyam,R.,
Strickland,C., Stilwell,J., Stone,C., Strickland,C.,

```

Subramanian,A., Sydney,K., Tang,L., Vassiliev,H., Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and Zody,M.  
 Direct Submission  
 Submitted (28-MAR-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 28, 1998 this sequence version replaced gi:2978476.  
 All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Only the last 48 kilobases of this clone are being submitted. The remainder overlaps accession number AC003070 (WICR project L266).

FEATURES  
 source  
 1. 48032  
 /location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_type="genomic DNA"  
 /mol\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clones="HRPCL169K15"  
 /clone\_lib="Peter de Jong/ human PAC library"  
 386..509  
 /rpt\_family="AluJo"  
 533..828  
 /rpt\_family="AluSx"  
 829..987  
 /rpt\_family="AluJo"  
 1211..1293  
 /rpt\_family="LINE2"  
 2194..2221  
 /rpt\_family="AT-rich"  
 complement(2222..2522)  
 /rpt\_family="AluSx"  
 3200..3413  
 /rpt\_family="MIR"  
 3809..4101  
 /rpt\_family="AluSx"  
 4103..4402  
 /rpt\_family="AluY"  
 complement(4493..4714)  
 /rpt\_family="MIR"  
 5205..5487  
 /rpt\_family="AluSx"  
 5617..5748  
 /rpt\_family="AluY"  
 complement(5933..6219)  
 /rpt\_family="AluSx"  
 6891..7191  
 /rpt\_family="AluSx"  
 7368..7666  
 /rpt\_family="AluSx"  
 complement(7638..7950)  
 /note="Single-stranded coverage."  
 complement(8021..8315)  
 /rpt\_family="AluSp"  
 complement(8316..8613)  
 /rpt\_family="AluJo"  
 complement(9398..9690)  
 /rpt\_family="AluSx"  
 complement(10095..10227)  
 /rpt\_family="FLAM\_C"  
 10680..10774  
 /rpt\_family="LINE2"  
 11480..11780  
 /rpt\_family="AluSx"  
 12408..12558  
 /rpt\_family="MIR"  
 12548..12607  
 /rpt\_family="LINE2"  
 complement(12969..13255)  
 /rpt\_family="AluJo"  
 14531..14691  
 /rpt\_family="MER3"



repeat_region	complement (14740..14818)	repeat_region	811	AGATCCAGCTCAGCCTTCAGCTTGGTCAACAAGACGGCAGCGCTGTTGGCTACGACACA	870
repeat_region	/rpt_family="MERSA"	repeat_region	44695	AGATCCAGCTCAGCCTTCAGCTTGGTCAACAAGACGGCAGCGCTGTTGGCTACGACACA	44754
repeat_region	/rpt_family="MERSA"	repeat_region	871	TGGAGGTGCCAGACTCGGGCATCGACCTCGAGTCACACTGGCCCTGATGGCAGCTTCG	930
repeat_region	16564..16662	repeat_region	44755	TGGAGGTGCCAGACTCGGGCATCGACCTCGAGTCACACTGGCCCTGATGGCAGCTTCG	44814
repeat_region	/rpt_family="MIR"	repeat_region	931	CCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCGCTTAACCTCCCTCCAC	990
repeat_region	/rpt_family="AluJb"	repeat_region	44815	CCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCGCTTAACCTCCCTCCAC	44874
repeat_region	18177..18205	repeat_region	991	CGCGGGCTCCACACTCGCGAAAGCAGCCTTCTGCTCGGTGCAGATGCTGCCCTGAAA	1050
repeat_region	/rpt_family="CA)n"	repeat_region	44875	CGCGGGCTCCACACTCGCGG--AAGCAGCCTTCTGCTCGGTGCAGATGCTGCCCTG-AA	44932
repeat_region	complement (19855..19953)	repeat_region	1051	ACACAGGCTCAGCCGTTCCACAGGGGATYTGNCACAGCCCCCGGCTCAGAGNTGGGAACC	1110
repeat_region	20824..21121	repeat_region	44933	ACACAGGCTCAGCCGTTCCACAGGG--ATCTGCGCAGCCCCCGGCTCAGCAGTGGGACCA	44989
repeat_region	/rpt_family="MIR"	repeat_region	1111	AGGCGCTCGNACAGNAGNAGGTTNGGGGCAAGCAGAGTGCCTCCAGGATTTCCACAN	1170
repeat_region	/rpt_family="AluSc"	repeat_region	44990	GGCCTCGCAGCAGCAAG-----GTGGGGGCAAGCAGAGTGCCTCCAGGATTTCCACAC	45042
repeat_region	complement (23516..23817)	repeat_region	1171	CCTGAGCCNTGCCCCCCTGCTGAADAAAAACATYVNCGCCACAGTGAAGACAGAGAAG	1230
repeat_region	/rpt_family="AluJo"	repeat_region	45043	C---TGAGCCCTGCCCCACCTGCTGAGAAAACT--CGGCCACGCTGAGAGACA---	45093
repeat_region	complement (23852..24153)	repeat_region	1231	GAGGATGNCAGGAGTTNNACCTYGGGAAACAAACAGGGATCTTTTCTGCCCTGC	1290
repeat_region	/rpt_family="AluSx"	repeat_region	45094	GAGGAGGATGGCAGGAGTTACCTCGGGAAACAAACAGGATCTTCTCTGCCCTGCTCCAG-	45152
repeat_region	complement (25772..25870)	repeat_region	1291	TCCAGTNGAGTTCGCTGCTGNACCGCTTGGANTCAGTGACCATTTGTTGGCAGANAGGG	1350
repeat_region	/rpt_family="MER46"	repeat_region	45153	-----TCGAGTTGGCTGGA---CCGCTTGGATCAGTGACCATTTGCTGGCAGA-CAGGG	45203
repeat_region	complement (26923..27203)	repeat_region	1351	GAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACCTTNC	1410
repeat_region	27261..27394	repeat_region	45204	GAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACCT-C	45262
repeat_region	/rpt_family="AluJo"	repeat_region	1411	CAGGCTGCTGAGAGAGTGCAAGTGTGAAGGNCACAAACCTCAGGNTTCAGTGAGA	1470
repeat_region	27554..27854	repeat_region	45263	CAGGCTGCTGAGAGTGTCAAGTGTGAAGGNCACAAACCTCAGGNTTCAGTGAGA	45318
repeat_region	/rpt_family="AluSgI"	repeat_region	1471	ACCAGGTNCAGCAGGATATGCCCGCCGNTAGGTTAANNNGGGGGCCCTCTNAAACCCCTTG	1530
repeat_region	27869..28156	repeat_region	45319	ACCAGGT-CAGCAGGATATGCCCGCCGNTAGGTTAANNNGGGGGCCCTCTNAAACCCCTTG	45373
repeat_region	28206..28382	repeat_region	1531	CTTNGGCCCTNACCTNGGCCAGCTCANCCCTTTTGGGTGTAGGGGAAAAAGATGCTTGA	1590
repeat_region	/rpt_family="AluSc"	repeat_region	45374	CCTGGCCCTCA---CCTGGCCAGCTCA-CCTCTTTTGGGTGTAGGGGAAAAAGATGCTTGA	45429
repeat_region	28587..28882	repeat_region	1591	CCCTGGGAAGGCTCCCTGGTGAAGATACACCACTTTTCAGGTTGTTGGCAACACAGGTC	1650
repeat_region	/rpt_family="AluSx"	repeat_region	45430	CCCTGGGAAGGCT-CCTGGTGAAGATACACCACTTTTCAGGTTGTTGGCAACACAGGTC	45488
repeat_region	28893..29195	repeat_region	1651	CTGAGTTGACCTCTGGTTTGGTTCAGCAAGGAGGTTGTAAGTGAAGTGGTTCTTCA	1710
repeat_region	/rpt_family="AluSx"	repeat_region	45489	CTGAGTTGACCTCTGGTTTGGTTCAGCAAGGAGGTTGTAAGTGAAGTGGTTCTTCA	45548
repeat_region	complement (29412..29506)	repeat_region	1711	GTNCCCAGACATGTGCCCTTTTGTGCTGGCTACCACTCTTCCCAGACAGCAGGCCCC	1770
repeat_region	/rpt_family="LINE2"	repeat_region	45549	GT-CCCCAGACATGTGCCCTTTTGTGCTGGCTACCACTCTTCCCAGACAGCAGGCCCC	45607
repeat_region	29775..30060				
repeat_region	/rpt_family="AluY"				
repeat_region	30071..30372				
repeat_region	/rpt_family="AluSp"				
repeat_region	30536..30817				
repeat_region	/rpt_family="AluJb"				
repeat_region	30859..31080				
repeat_region	/rpt_family="MIR"				
repeat_region	complement (31106..31144)				
repeat_region	/rpt_family="MIR"				
repeat_region	32189..32335				
repeat_region	/rpt_family="L1PA4"				
repeat_region	32370..32670				
repeat_region	/rpt_family="AluSg"				
repeat_region	32685..32979				
repeat_region	/rpt_family="AluSx"				
repeat_region	33437..33738				
repeat_region	/rpt_family="AluSg"				
repeat_region	complement (33995..34315)				
repeat_region	/rpt_family="AluSc"				
repeat_region	complement (34339..34639)				
repeat_region	/rpt_family="AluY"				
repeat_region	complement (36539..36839)				
repeat_region	/rpt_family="AluSx"				
repeat_region	complement (37152..37449)				
repeat_region	/rpt_family="AluJo"				
repeat_region	37854..38155				
repeat_region	/rpt_family="AluSx"				
repeat_region	39172..39237				
repeat_region	/rpt_family="MIR"				
repeat_region	39346..40104				
repeat_region	/rpt_family="MER21B"				
repeat_region	40186..40475				
repeat_region	/rpt_family="AluJo"				
repeat_region	40946..41076				

repeat\_region  
ORIGIN  
Query Match  
Best Local Similarity  
Matches 1487; Conservative

37.0%; Score 973.4; DB 9; Length 48032;  
84.4%; Pred. No. 1.3e-238;  
5; Mismatches 183; Indels 87; Gaps 35;

/rpt\_family="FLAM\_A"  
43760..44054  
/rpt\_family="AluSx"





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:57:41 ; Search time 30.0772 Seconds  
(without alignments)  
2390.553 Million cell updates/sec

Title: US-09-155-676B-2  
Perfect score: 3093  
Sequence: 1 XTGPGXGMSGXNXXKQF.....FLMTFHVCEPINCFSLSLKKK 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 557933 seqs, 119041758 residues

Total number of hits satisfying chosen parameters: 557933

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:\*

1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US11\_NEW\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3058	98.9	604	5	US-09-155-676B-2
2	388	12.5	91	6	US-10-450-763-30582
3	302	9.8	520	7	US-11-097-143-11289
4	236	7.6	131	6	US-10-450-763-58372
5	193.5	6.3	1082	7	US-11-090-997-344
6	189	6.1	1094	7	US-11-090-997-1432
7	183.5	5.9	720	7	US-11-097-143-14316
8	181.5	5.9	269	8	US-60-655-875-154870
9	175.5	5.7	473	1	PCT-US03-10870-2491
10	171	5.5	1006	6	US-10-450-763-51537
11	160	5.2	366	8	US-60-669-175-21377
12	154	5.0	388	6	US-10-450-763-51533
13	151	4.9	600	8	US-60-680-002-931
14	151	4.9	601	8	US-60-671-501-404
15	143	4.6	183	6	US-10-450-763-51534
16	139.5	4.5	273	8	US-60-669-241-48091
17	136.5	4.4	2844	1	PCT-US03-10870-1970
18	136.5	4.4	2844	8	US-60-671-501-211
19	136	4.4	1139	6	US-10-940-774A-6556
20	136	4.4	1146	6	US-10-940-774A-7567
21	130.5	4.2	440	6	US-10-940-774A-9774
22	130	4.2	2829	6	US-10-525-621-1
23	123	4.0	972	8	US-60-643-717-15402
24	122	3.9	1945	7	US-11-097-143-21633
25	121.5	3.9	1035	7	US-11-097-143-2004

## ALIGNMENTS

## RESULT 1

US-09-155-676B-2  
; Sequence 2, Application US/09155676B  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR  
; FILE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: WALLACH-21  
; CURRENT APPLICATION NUMBER: US/09/155.676B  
; CURRENT FILING DATE: 1999-01-04  
; PRIOR APPLICATION NUMBER: PCT/IL97/00117  
; PRIOR FILING DATE: 1997-04-01  
; PRIOR APPLICATION NUMBER: IL 117800  
; PRIOR FILING DATE: 1996-04-02  
; PRIOR APPLICATION NUMBER: IL 119133  
; PRIOR FILING DATE: 1996-08-26  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 2  
; LENGTH: 604  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (6)..(6)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (8)..(8)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (13)..(13)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (15)..(15)  
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
; FEATURE:  
; NAME/KEY: misc feature

Sequence 1425, Ap  
Sequence 10758, A  
Sequence 10759, A  
Sequence 2013, Ap  
Sequence 45794, A  
Sequence 54021, A  
Sequence 34486, A  
Sequence 1186, Ap  
Sequence 29559, A  
Sequence 1655, Ap  
Sequence 12774, A  
Sequence 23, Appl  
Sequence 23, Appl  
Sequence 10757, A  
Sequence 42824, A  
Sequence 47387, A  
Sequence 1586, Ap  
Sequence 943, App  
Sequence 943, App  
Sequence 943, App

26 120.5 3.9 2641 1 PCT-US03-10870-1425  
27 119.5 3.9 1211 6 US-10-990-328A-10758  
28 121.5 3.9 1211 6 US-10-990-328A-10759  
29 119.5 3.9 1528 7 US-11-097-143-2013  
30 118.5 3.8 904 6 US-10-450-763-45794  
31 118.5 3.8 904 6 US-10-450-763-54021  
32 118.5 3.8 1015 6 US-10-450-763-34486  
33 118 3.8 1478 6 US-10-489-448-1186  
34 117 3.8 1980 7 US-11-097-143-29559  
35 116 3.8 1042 1 PCT-US03-10870-1655  
36 116 3.8 1351 7 US-11-097-143-12774  
37 115.5 3.7 490 7 US-11-063-325-23  
38 115.5 3.7 490 7 US-11-011-526A-23  
39 114.5 3.7 1218 6 US-10-990-328A-10757  
40 113 3.7 655 6 US-10-450-763-42824  
41 112.5 3.6 195 6 US-10-450-763-47387  
42 112.5 3.6 487 7 US-11-090-997-1586  
43 112.5 3.6 516 1 PCT-US04-17965-943  
44 112.5 3.6 516 1 PCT-US04-17965B-943  
45 112.5 3.6 516 1 PCT-US04-17965C-943

LOCATION: (37)..(37)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (271)..(271)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (274)..(274)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (334)..(334)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (348)..(348)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (354)..(355)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (359)..(359)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (363)..(363)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (405)..(405)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (549)..(549)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (569)..(570)  
OTHER INFORMATION: Xaa can be any naturally occurring amino acid

US-09-155-676B-2

Query Match 98.9%; Score 3058; DB 5; Length 604;  
Best Local Similarity 100.0%; Pred. No. 5.2e-221; Indels 0; Gaps 0;  
Matches 603; Conservative 0; Mismatches 0;

QY 2 TGPXGXMGSXNKKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCCLCAQFEAVL 61  
DB 2 TGPXGXMGSXNKKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCCLCAQFEAVL 61  
QY 62 QHGLKRSRGLATAAAIKQAAGFASKTETEPVFWYVYVKEVNLKHELORFYSLRHIAADVG 121  
DB 62 QHGLKRSRGLATAAAIKQAAGFASKTETEPVFWYVYVKEVNLKHELORFYSLRHIAADVG 121  
QY 122 RGAWLRCALNEHSLERLYLHMLADRCRLSTFYEDWSFVMDERRSSMLPTMAAGLSILF 181  
DB 122 RGAWLRCALNEHSLERLYLHMLADRCRLSTFYEDWSFVMDERRSSMLPTMAAGLSILF 181  
QY 182 AINIDNDKLNQOSKFAFTVSDLLKSTQNTVTSLLKSTQGVSSLFREITASSAVSILIKP 241  
DB 182 AINIDNDKLNQOSKFAFTVSDLLKSTQNTVTSLLKSTQGVSSLFREITASSAVSILIKP 241  
QY 242 EQETDPCLSCPGMSVLPNNAKRSRRKRPXTXSHLMRKMSRTLTCLKSHLGGQRAOR 301  
DB 242 EQETDPCLSCPGMSVLPNNAKRSRRKRPXTXSHLMRKMSRTLTCLKSHLGGQRAOR 301  
QY 302 TTPTAPLISICPPILKAPSGLTPMSEQOLMENXFPVFERGVWVPEAXCEKHXRCXKX 361  
DB 302 TTPTAPLISICPPILKAPSGLTPMSEQOLMENXFPVFERGVWVPEAXCEKHXRCXKX 361  
QY 362 RXRVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAEGAAQXLRHPLPCQWRGLLOPS 421

DB 362 RXRVWKLIRKEAQGLGVAREATGREHLPLPDAQLGSAEGAAQXLRHPLPCQWRGLLOPS 421  
QY 422 RCPRPKGERDRTRGPRSPGTSVOCGSQLSRPKSSEQPVTSASVPESMTISELRQAT 481  
DB 422 RCPRPKGERDRTRGPRSPGTSVOCGSQLSRPKSSEQPVTSASVPESMTISELRQAT 481  
QY 482 VAMNRKDELEENRSIRNLLDGEHSAALROEVDTLKRKVAEQERQGMKVQALASYL 541  
DB 482 VAMNRKDELEENRSIRNLLDGEHSAALROEVDTLKRKVAEQERQGMKVQALASYL 541  
QY 542 CYFVRFPXPHVRTMMWRNGREKSNSSXXSHLSSWTQSFILKCLFMTFHVCEPINCFSHL 601  
DB 542 CYFVRFPXPHVRTMMWRNGREKSNSSXXSHLSSWTQSFILKCLFMTFHVCEPINCFSHL 601  
QY 602 KKK 604  
DB 602 KKK 604

RESULT 2

US-10-450-763-30582  
; Sequence 30582, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 30582  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-450-763-30582

Query Match 12.5%; Score 388; DB 6; Length 91;  
Best Local Similarity 96.3%; Pred. No. 7.4e-22;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 MSGSNXNDKROFLERLLDAVKQCQIRFXGRKEIASDSRVTCCLCAQFEAVLQHLKRS 68  
DB 1 MSGSNXNDKROFLERLLDAVKQCQIRFXGRKEIASDSRVTCCLCAQFEAVLQHLKRS 60  
QY 69 RGLALTAATAAIKQAAGFASKTET 90  
DB 61 RGLALTAATAAIKQAAGFASKTET 82

RESULT 3

US-11-097-143-11289  
; Sequence 11289, Application US/11097143  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932





;; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF  
;; FILE REFERENCE: 38-21(53885)  
;; CURRENT APPLICATION NUMBER: US/60/655.875  
;; CURRENT FILING DATE: 2005-02-24  
;; NUMBER OF SEQ ID NOS: 171306  
;; SEQ ID NO 154870  
;; LENGTH: 269  
;; TYPE: PRT  
;; ORGANISM: Heterodera glycines

;; NAME/KEY: misc feature  
;; LOCATION: (241) (241)  
;; OTHER INFORMATION: Xaa can be any naturally occurring amino acid

;; FEATURE:  
;; OTHER INFORMATION: Coding regions on vcdna: vcdna-SeqID 81293; Strands=-; Position=1  
;; OTHER INFORMATION: -166,1016-1128,1249-1344,1396-1475,1529-1639,1721-1805,3272-3393  
;; FEATURE:  
;; OTHER INFORMATION: Homolog annotation: Hit\_ID=NP\_496906.2; Match level="QueryCovered  
;; OTHER INFORMATION: =96%, HitCoverage=41%, E-value=2e-49, Identity=41%; Hit descrip  
;; OTHER INFORMATION: =RUN domain containing protein (20146) [Caenorhabditis elegans]  
;; OTHER INFORMATION: emb|CAA21666.2| Hypothetical protein Y51H1  
US-60-655-875-154870

Query Match 5.9%; Score 181.5; DB 8; Length 269;  
Best Local Similarity 26.9%; Pred. No. 8.9e-06;  
Matches 60; Conservative 33; Mismatches 81; Indels 49; Gaps 8;

QY 27 DAVKQCOIR-----FXGRKIASDSDSRVT-CLCAQFEAVLQHLKRSRGLATAAAI 78  
DB 32 DVEKKCKERTESFLKTLGSRK---SQISSELTQALCTVIEAVFHIGURDAFFL----- 81  
QY 79 KQAGFASKTETEPVFWYVYKVLNKHLEQFYSLRHIAASDVGRGR-----AMLRCA 130  
DB 82 ---KGSRRSNRPSNFWPFWKYQTSIKQISLNLQIHTEIGRARAWIRVLNANRIV 138  
QY 131 LNEHSLERYLHMLADRCRL-----STRYEDWSFVMDERSMPLPTMAAGLNSILFAINID 186  
DB 139 LNEYSLDHYISMFKDTROMKQFYAQFYADDAFLRDS-----CVNSLGVH 184  
QY 187 NKDLNGOSKEAPTVDLLKESTQNTVSL--LKESTQGVSSLER 227  
DB 185 LKRINGLKITAPTNSLSLNTWTPSLSLAGLIKTAPSIGGAFK 227

RESULT 9

PCT-US03-10870-2491  
; Sequence 2491, Application PC/TUS0310870

;; GENERAL INFORMATION:  
;; APPLICANT: Mitokor, Inc.  
;; APPLICANT: Buck Institute  
;; APPLICANT: Ghosh, Soumitra S.  
;; APPLICANT: Fahy, Eoin D.  
;; APPLICANT: Zhang, Bing  
;; APPLICANT: Gibson, Bradford W.  
;; APPLICANT: Taylor, Steven W.  
;; APPLICANT: Glenn, Gary M.  
;; APPLICANT: Warnock, Dale E.

;; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
;; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
;; FILE REFERENCE: 660088.465PC  
;; CURRENT APPLICATION NUMBER: PCT/US03/10870  
;; CURRENT FILING DATE: 2003-04-04  
;; NUMBER OF SEQ ID NOS: 3025

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 2491

;; LENGTH: 473

;; TYPE: PRT

;; ORGANISM: Homo sapiens

PCT-US03-10870-2491

Query Match 5.7%; Score 175.5; DB 1; Length 473;  
Best Local Similarity 28.5%; Pred. No. 5.2e-05;  
Matches 75; Conservative 35; Mismatches 118; Indels 35; Gaps 12;

QY 5 GXGMSGXNXXKROFLLERLLDAVKQCQIRFXGRK-----EIASDSDSRVTCCLAQF 57  
DB 14 GGGGGGKKSLSARNAVER-RNLITVC--RFSVKTLIDRSCTETIDDSPEFNNFAAIL 70  
QY 58 EAVLQHLKR----SRGLATAAAIK-QAAGFASKTETEPVFWYVYKVLN-----HELQR 109  
DB 71 EQILSHRLKEISQSCRWLAHLQIPLOGQVTFG--YESPRSFMDYIRVACRKVSQNCICS 128  
QY 110 FYSLRHIASDVGRGRAWLRCAALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSM 169  
DB 129 IENWENVSSRAKRAWIRVALMEKHLSEYISTALRDFKTRRFYEDGAIVLGE--ANML 187  
QY 170 PTMAAGLNSILFAINIDNKLNGOSKFAPTVSDLLKESTQNTVSLK--ESTQGVSSLF 227  
DB 188 AGMLGLNAIDFSCFKGEGLDGS---FPAVID-----YTPYLKVIQSSDISDEE 236  
QY 228 EI-TASSAVSILIKPQETDPCL 249  
DB 237 ELRTLGGSGSESSTPENVGPPFL 259

RESULT 10

US-10-450-763-51537  
; Sequence 51537, Application US/10450763

;; GENERAL INFORMATION:  
;; APPLICANT: Hyseq, Inc  
;; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
;; FILE REFERENCE: 790CIP3/US  
;; CURRENT APPLICATION NUMBER: US/10/450,763  
;; CURRENT FILING DATE: 2003-06-11  
;; PRIOR APPLICATION NUMBER: PCT/US01/08631  
;; PRIOR FILING DATE: 2001-03-30  
;; PRIOR APPLICATION NUMBER: 09/540,217  
;; PRIOR FILING DATE: 2000-03-31  
;; PRIOR APPLICATION NUMBER: 09/649,167  
;; PRIOR FILING DATE: 2000-08-23  
;; NUMBER OF SEQ ID NOS: 60736  
;; SOFTWARE: Custom  
;; SEQ ID NO 51537  
;; LENGTH: 1006  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: DOMAIN

;; LOCATION: (957) (973)

;; OTHER INFORMATION: Phorbol esters / diacylglycerol binding domain proteins  
;; OTHER INFORMATION: domain identified by eMATRIX, accession number BL004798, p-value  
;; OTHER INFORMATION: 7.429e-09, raw score of 12.57

;; FEATURE:

;; NAME/KEY: DOMAIN

;; LOCATION: (476) (721)

;; OTHER INFORMATION: PH domain identified by PFam, accession name PH, E-value=2.5e  
;; OTHER INFORMATION: -09, PFam score of 38.4

;; FEATURE:

;; NAME/KEY: misc feature

;; LOCATION: (1) (1006)

;; OTHER INFORMATION: Xaa = X or \* as defined in Table 2

US-10-450-763-51537

Query Match 5.5%; Score 171; DB 6; Length 1006;

Best Local Similarity 22.7%; Pred. No. 0.00029;

Matches 103; Conservative 55; Mismatches 157; Indels 138; Gaps 20;

QY 83 GFASKTETEPVFWYVYKVLNKHLEQFYSLRHIAASDVGRGRAWLRCAALNEHSLERYLHM 142

DB 19 GFSSPAVQEANRGALCSSIERGHIISELEHLTFVNTDVGRCAWLRLALNDGLMECYLKL 78

QY 143 LLADRCRLSTFYEDWSFVMDERSMPLPTMAAGLNSILFAINIDNKDLN-----CQSK 195

DB 79 LLQEQARLHEYYOPTALLRDAEAGEFLLSFLOGLTSLSFELSYKSAILNEWTLTPLALSG 138

QY 196 FAPTVDLLKESTQNTVSLLKESTQGVSSLF-----REITASS----- 233

```
Db 139 LCP-LSELDPLSTGAELQKESLDSISHSSGSDIEVHSHGKIRRNQKLTASSLSLDT 197
QY 234 AVSILIKPEQETDPCL-----SCPGMSVLMPNNAKSGRRKXPTXSHL 277
Db 198 ASSQSCLNSDCLQENGSKSPDHCEPMSDSD-DLGTANAEDSDSLQE----- 249
QY 278 MMRKMSRTLTGTLKRLHGQGAQRT---TPTAPLS-----TSCPLKAPSGLTME 325
Db 250 VLLFESKAQVNSVPTN---GLSQETETPTQASLSLHGLNTYTLHC---EAPAEPLAQ 303
QY 326 SEQOLMENKFPFVERGVWPEAXCEKHXRXRCXKRXRWKLIIRKEAGPGVAR--BA 383
Db 304 AASGTQD-----GVHVQEP-----RPAQSPDLQOPVES 333
QY 384 TGREHLPLPDA-----OLGSAEGAAOXLRHPLPCQWRGLLQPSRCPRKPGERDRTG 436
Db 334 TSGQQ---PSSTVSETAREVGQGNLQKAQAD-----GAGLKLIVSS 373
QY 437 PRSP--GSWTSVQCGSQLSRPRKSSQPVTSAS 467
Db 374 PTSPKMKMIS---EDFEYRP--SREQPLESAS 401

RESULT 11
US-60-669-175-21377
; Sequence 21377, Application US/60669175
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
; APPLICANT: Baum, James A
; APPLICANT: Gilbertson, Larry A
; APPLICANT: Kovalic, David K
; APPLICANT: LaRosa, Thomas J
; APPLICANT: Lu, Maolong
; APPLICANT: Munvikwa, Tichifa R. I.
; APPLICANT: Roberts, James K
; APPLICANT: Wu, Wei
; APPLICANT: Zhang, Bei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CONTROL OF INSECT INFESTATION IN PLANTS
; FILE REFERENCE: 38-21(53597)
; CURRENT APPLICATION NUMBER: US/60/669,175
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: 60560842
; PRIOR FILING DATE: 2004-04-09
; PRIOR APPLICATION NUMBER: 60565632
; PRIOR FILING DATE: 2004-04-27
; PRIOR APPLICATION NUMBER: 60579062
; PRIOR FILING DATE: 2004-06-11
; PRIOR APPLICATION NUMBER: 60603421
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60617261
; PRIOR FILING DATE: 2004-10-11
; NUMBER OF SEQ ID NOS: 40774
; SEQ ID NO 21377
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Diabrotica virgifera
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_1074; Strand=+; Position=317
; OTHER INFORMATION: Homolog annotation: Hit_ID=XP_395844.1; Match level="QueryCoverage=99%, HitCoverage=60%, E-value=4e-64, Identity=41%"; Hit description: "similar to CG6613-PA [Apis mellifera]"
; OTHER INFORMATION: Pfam annotation: Pfam ID=RUN; Match level="Score=108.7, E-value=1e-29, Copies=1"; Pfam description=RUN domain
US-60-669-175-21377

Query Match 5.2%; Score 160; DB 8; Length 366;
Best Local Similarity 23.0%; Pred. No. 0.00054;
Matches 62; Conservative 55; Mismatches 105; Indels 48; Gaps 9;
QY 18 RQFLERLLDAVKQCIQIRFXKKEATASDSDR-----VTCLCAQFEAVLQHLKRSGL 71
```

```
Db 19 KKSITRQLSDCK--EIHVAGVEENKSPNFSEFSTTAATLCTAIEAIFLHGLRDT--- 73-
QY 72 ALTRAAAIKQAAGFASKTETEPVFWYVYKVLNKHLELQRFYSLRHIASDVGRGAWLRAL 131
Db 74 --LIHKFKALADVE-QSEPSFWAPLLIISHRQIIIEQITNLSQITTEVGQCRVIRLAL 130
QY 132 NEHSILERYLHMLLADRCLSTFYEDWSFVMDERSSMLPTMAAGLNSILFAINDKDLN 191
Db 131 NDCULSSYLMIRQDSSSLKSYKYKNAYVRDGLLDVAORLIEGVEAF----- 178
QY 192 GQSKFAPTVSDLLKESTQNTVSLKKESTQGVSS-----LFRITASSAVSIL 238
Db 179 -KSFTLPNSSL--NTWPLPSLI---LSGLWSFTLSCPVAPCDVQAQSIITESMKYQ 232
QY 239 IKPQETDPCLSGCMGMSVLMPNNAKSGRRK 268
Db 233 SNFENSSDTASIC---SYMVSNSQSSGLRQ 259

RESULT 12
US-10-450-763-51533
; Sequence 51533, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51533
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-51533

Query Match 5.0%; Score 154; DB 6; Length 388;
Best Local Similarity 25.7%; Pred. No. 0.0016;
Matches 53; Conservative 29; Mismatches 76; Indels 48; Gaps 6;
QY 95 WY----YVKEVNLKHELQRFYSLRHIASDVGRGAWLRALNEHSILERYLHMLLADRCL 151
Db 53 WWRMDWTPLRPSRHIIELEHLTFVNMVGRCAWLRLALNGLMECYLKLLEQEARLR 112
QY 152 TFYEDWSFVMDERSSMLPTMAAGLNSILFAINDKDLN-----GQSKFAPTVSDLL 204
Db 113 EYOPTALLRDABEGEFLSLFLOGLMSLSPELSYKSAILNEWTLTPLALSGLCP-LSELD 171
QY 205 KESTQNTVSLKKESTQGV-----SSIFRITASSAVSIL 238
Db 172 PLSTSGAELQKESLDSISHSSGSDIEVHSHGKIRRNQKPTASSLSLDTASSQLSCS 231
QY 239 IKPQETDPCLSGCMGMSVLMPNNAKRS 264
Db 232 L-----NSDSCL-----LOENGSKS 246

RESULT 13
US-60-680-002-931
; Sequence 931, Application US/60680002
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001601
; CURRENT APPLICATION NUMBER: US/60/680,002
; CURRENT FILING DATE: 2005-05-12
```



```
; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 931
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-931

Query Match
  4.9%; Score 151; DB 8; Length 600;
Best Local Similarity 21.3%; Pred. No. 0.0048;
Matches 111; Conservative 71; Mismatches 190; Indels 150; Gaps 20;

Qy 56 QFEAVLQHLKRSGLALTAATAAIKQAAGFASKTETEPVFWYVYKVLNKLHQLQRFYSLRH 115
Db 39 QFFVMEHCLK--HGLKVKSFICQNKSPFGPLELVEKLCPEASDIATS-----VRNLPE 91
Qy 116 IASDVGRGRAWLRCAALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSSMLPTMAAG 175
Db 92 LKTAGRGRAWLYALMOKKLADYLVKIDLNKHLSEFYEPALMMEEE--GMVIVGLLVG 150
Qy 176 LNSILFAINDNKLNGQ-----SKFAPTVSDI--LKESTQNTVTSLLKESTQGVSSLFR 227
Db 151 LNVLDANLCLKGEDLDQSGVIDFSLYLKDVQDLGGKEH-ERITVDLDQKNY-----202
Qy 228 EITASSAVSILIRPEQETDPCLCSPGMSVLPNAPKRSRRKRXKPTXSHLMRKMSTLGL 287
Db 203 -----VEELNRHLSC---TVGDLQTKIDGLEK-----TNSKLQEELSAATDR 241
Qy 288 TCLKRHLGQGAQRTTTPAPLISCPPLKAPSGLTPMSEQOLMENXFPVFERGVWVPE- 346
Db 242 IC-----SLOEQOQLREQNELIRERSEKSEVEI 269
Qy 347 -----AXCEKHXRXRG--KRRXRVMKLRKEAQGPGLVAREATGREHLPLPDAQLG-- 397
Db 270 TKQDTKVELEYTKTQGLDEMYSVMKQKLEKKEKVRLELEKEL-----ELQIGMK 320
Qy 398 -SAEGAAQXLRHLPQWGRGLLPSCRPPKPRGDRTRGPRSPGWSVTSVQCGSOLSRPR 456
Db 321 TEMEIAMKLL-----EKD-THEKQD-----TLVALRQOLEBEVK 352
Qy 457 KSSQOPVTSASVPESM-----TISELRQATVAMNRKDELEENRSLRNLLDGEHSA 510
Db 353 AINLQMFHKAQNAESSLOQKNEAITSFEGKTNQVMSMKQMEERLOHSEARARQGAERSH 412
Qy 511 ALRQEV-----DTLKRKV-AEQEERQGMK 533
Db 413 KLOQELGGRIGALQQLSOLHEQCSSLKELKSEKEQORALQ 454

RESULT 14
US-60-671-501-404
; Sequence 404, Application US/60671501
; GENERAL INFORMATION:
; APPLICANT: Wim Van Criekinge
; TITLE OF INVENTION: METHYLATION MARKERS FOR DIAGNOSIS AND
; FILE REFERENCE: 006791.00001
; CURRENT APPLICATION NUMBER: US/60/671.501
; CURRENT FILING DATE: 2005-04-15
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 404
; LENGTH: 601
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-671-501-404

Query Match
  4.9%; Score 151; DB 8; Length 601;
Best Local Similarity 21.3%; Pred. No. 0.0048;
Matches 111; Conservative 71; Mismatches 190; Indels 150; Gaps 20;

Qy 56 QFEAVLQHLKRSGLALTAATAAIKQAAGFASKTETEPVFWYVYKVLNKLHQLQRFYSLRH 115
Db 39 QFFVMEHCLK--HGLKVKSFICQNKSPFGPLELVEKLCPEASDIATS-----VRNLPE 91
Qy 116 IASDVGRGRAWLRCAALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSSMLPTMAAG 175
Db 92 LKTAGRGRAWLYALMOKKLADYLVKIDLNKHLSEFYEPALMMEEE--GMVIVGLLVG 150
Qy 176 LNSILFAINDNKLNGQ-----SKFAPTVSDI--LKESTQNTVTSLLKESTQGVSSLFR 227
Db 151 LNVLDANLCLKGEDLDQSGVIDFSLYLKDVQDLGGKEH-ERITVDLDQKNY-----202
Qy 228 EITASSAVSILIRPEQETDPCLCSPGMSVLPNAPKRSRRKRXKPTXSHLMRKMSTLGL 287
Db 203 -----VEELNRHLSC---TVGDLQTKIDGLEK-----TNSKLQEELSAATDR 241
Qy 288 TCLKRHLGQGAQRTTTPAPLISCPPLKAPSGLTPMSEQOLMENXFPVFERGVWVPE- 346
Db 242 IC-----SLOEQOQLREQNELIRERSEKSEVEI 269
Qy 347 -----AXCEKHXRXRG--KRRXRVMKLRKEAQGPGLVAREATGREHLPLPDAQLG-- 397
Db 270 TKQDTKVELEYTKTQGLDEMYSVMKQKLEKKEKVRLELEKEL-----ELQIGMK 320
Qy 398 -SAEGAAQXLRHLPQWGRGLLPSCRPPKPRGDRTRGPRSPGWSVTSVQCGSOLSRPR 456
Db 321 TEMEIAMKLL-----EKD-THEKQD-----TLVALRQOLEBEVK 352
Qy 457 KSSQOPVTSASVPESM-----TISELRQATVAMNRKDELEENRSLRNLLDGEHSA 510
Db 353 AINLQMFHKAQNAESSLOQKNEAITSFEGKTNQVMSMKQMEERLOHSEARARQGAERSH 412
Qy 511 ALRQEV-----DTLKRKV-AEQEERQGMK 533
Db 413 KLOQELGGRIGALQQLSOLHEQCSSLKELKSEKEQORALQ 454

RESULT 15
US-10-450-763-51534
; Sequence 51534, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 51534
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-51534

Query Match
  4.6%; Score 143; DB 6; Length 183;
Best Local Similarity 28.6%; Pred. No. 0.0042;
Matches 53; Conservative 24; Mismatches 80; Indels 28; Gaps 5;

Qy 54 CAQFEAVLQHLKRSGLALTAATAAIKQAAGFASKTETEPVFWYVYKVLNKLHQLQRFYSL 113
Db 20 CTPSTSKLPQGGK-----KKSA--HOKPLPQPVFWPLLKATVTRKHIISEHL 66
Qy 114 RHIASDVGRGRAWLRCAALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSSMLPTMA 173
Db 67 TFVMDVGRGRAWLRALNGLMECYLKLQEQARLQ-----QRKESLOSIS 115
Qy 174 AGLNSILFAINDNKLNGQSGFAPTVSDI--LKESTQNTVTSLLKESTQGVSSLFRITASS 233
```

Db 116 HSSGSEDIEVHHGHKIRNQK--PTASSLSLDTAS--SSQLSCSLNCDSCLLQENGSKS 171  
QY 234 AVSIL 238  
Db :|  
172 PDHVL 176

Search completed: June 7, 2005, 12:21:16  
Job time : 31.0772 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:52:21 ; Search time 269.467 Seconds  
(without alignments)  
2618.037 Million cell updates/sec

Title: US-09-155-676B-2

Perfect score: 3093

Sequence: 1 XTGPXGXMGSGXNXDKRQF.....FLMTFHVCEPINCFSHLKKX 604

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCRUS\_COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pcp.\*
- 33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pcp.\*
- 34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pcp.\*
- 35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pcp.\*
- 36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pcp.\*
- 37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	98.9	3058	604	15	US-09-155-676-2 Sequence 2, Appli
2	98.9	3058	604	15	US-09-155-676A-2 Sequence 2, Appli
3	1594.5	51.6	541	26	US-10-070-255-7 Sequence 7, Appli
4	1537.5	49.7	784	26	US-10-070-255-9 Sequence 9, Appli
5	1537.5	49.7	813	26	US-10-070-255-8 Sequence 8, Appli
6	1196	38.7	375	20	US-09-629-469A-14017 Sequence 14017, A
7	1196	38.7	375	35	US-10-917-503-14017 Sequence 14017, A
8	1194	38.6	250	27	US-10-170-205B-18638 Sequence 18638, A
9	721	23.3	320	22	US-09-758-472-5854 Sequence 5854, Ap
10	721	23.3	320	28	US-10-235-926-5854 Sequence 5854, Ap
11	693.5	22.4	548	37	US-60-212-356-236 Sequence 236, App
12	685.5	22.2	544	37	US-60-212-356-203 Sequence 203, App
13	663.5	21.5	692	37	US-60-229-525-294 Sequence 294, App
14	498.5	16.1	116	27	US-10-170-205B-22102 Sequence 22102, A
15	388	12.5	91	1	PCT-US01-08631-30582 Sequence 30582, A
16	374.5	12.1	424	27	US-10-170-205B-6178 Sequence 6178, Ap
17	348.5	11.3	329	37	US-60-229-525-456 Sequence 456, App
18	311	10.1	411	27	US-10-170-205B-6243 Sequence 6243, Ap
19	302	9.8	520	20	US-09-614-150-11289 Sequence 11289, A
20	302	9.8	520	20	US-09-614-150A-11289 Sequence 11289, A
21	302	9.8	520	37	US-60-167-217-11318 Sequence 11318, A
22	302	9.8	520	37	US-60-173-464-9160 Sequence 9160, Ap
23	302	9.8	520	37	US-60-191-637-11321 Sequence 11321, A
24	302	9.8	520	37	US-60-191-681-8867 Sequence 8867, Ap
25	242.5	7.8	75	27	US-10-170-205B-17822 Sequence 17822, A
26	236	7.6	131	1	PCT-US01-08631-58372 Sequence 58372, A
27	199.5	6.5	522	21	US-09-724-676-66567 Sequence 66567, A
28	199.5	6.5	522	21	US-09-724-676A-66567 Sequence 66567, A
29	183.5	5.9	720	20	US-09-614-150-14316 Sequence 14316, A
30	183.5	5.9	720	20	US-09-614-150A-14316 Sequence 14316, A
31	183.5	5.9	720	37	US-60-167-217-14393 Sequence 14393, A
32	183.5	5.9	720	37	US-60-191-637-14357 Sequence 14357, A
33	183	5.9	673	21	US-09-724-676-66566 Sequence 66566, A
34	183	5.9	673	21	US-09-724-676A-66566 Sequence 66566, A
35	183	5.9	788	21	US-09-724-676-66557 Sequence 66557, A
36	183	5.9	788	21	US-09-724-676-66558 Sequence 66558, A
37	183	5.9	788	21	US-09-724-676-66559 Sequence 66559, A
38	183	5.9	788	21	US-09-724-676-66560 Sequence 66560, A
39	183	5.9	788	21	US-09-724-676-66561 Sequence 66561, A
40	183	5.9	788	21	US-09-724-676-66562 Sequence 66562, A
41	183	5.9	788	21	US-09-724-676-66563 Sequence 66563, A
42	183	5.9	788	21	US-09-724-676-66564 Sequence 66564, A
43	183	5.9	788	21	US-09-724-676-66565 Sequence 66565, A
44	183	5.9	788	21	US-09-724-676A-66557 Sequence 66557, A
45	183	5.9	788	21	US-09-724-676A-66558 Sequence 66558, A

ALIGNMENTS

RESULT 1  
US-09-155-676-2  
; Sequence 2, Application US/09155676  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P L L C  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001





Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLSILFAINIDNK 180  
Qy 189 DLNGOSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREITASSAVSILIKPEQETDP 248  
Db 181 DLNGOSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREITASSAVSILIKPEQETDP- 239  
Qy 249 LSCPGMSVLMPN-----AKRSGRKRKXPTXSHLMRKMRTLTGCLKRLHGGQRAQR- 301  
Db 240 -----LPVVRNVSADAKCKKXKVTNIIISFDDDEEQNSGDVFKKTPGAGESSED 294  
Qy 302 -----TTPTAPLS-----ISCPLKAPSGLTPME-----SEQOOLMENX 334  
Db 295 NSDRSSVNMISAFEPFGPNSSQSSNSWKIDSLSLNGEFGYQKLDVKSIDDDVDENE 354  
Qy 335 FVPERGVWVPEACEKX-----XRCCKKRRXRVWKLIRKEAQGLGVARE- 382  
Db 355 DDVY-----GNSGRKRGHSESPKPLEGNTC-----LSQMSWAPLKVHLND 398  
Qy 383 -----ATGREHLPLPDAQLSAEGAAQXLRHPLPCQWRGLLPQSRCPKPRGDRDRTRG 436  
Db 399 SDILFPVSGVGSYPADAPLGSLNGTGPEDHVL-----DPGLRYSVEA 443  
Qy 437 PRSPGWSVQCGSOLSRPKSSQPVTSASVPESMTISELRQATVAMNRRKDELEENR 496  
Db 444 -SSPGH-----GSPLS-----SLLPSASVPESMTISELRQATVAMNRRKDELEENR 489  
Qy 497 SLRNLGEMHSAAALROEVDTLKRVABOERQGMKVOALA 538  
Db 490 SLRNLGEMHSAAALROEVDTLKRVABOERQGMKVOALA 531

## RESULT 5

US-10-070-255-8  
; Sequence 8, Application US/10070255  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: WALLACH, Nikolay  
; APPLICANT: SINHA, Indranil  
; APPLICANT: LEU, Stefan  
; TITLE OF INVENTION: IREN PROTEIN, ITS PREPARATION AND USE  
; FILE REFERENCE: WALLACH=28  
; CURRENT APPLICATION NUMBER: US/10/070,255  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/IL00/00517  
; PRIOR FILING DATE: 2000-08-31  
; PRIOR APPLICATION NUMBER: IL 131719  
; PRIOR FILING DATE: 1999-09-02  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 813  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-070-255-8

Query Match 49.7%; Score 1537.5; DB 26; Length 813;  
Best Local Similarity 61.2%; Pred. No. 4.2e-125;  
Matches 356; Conservative 22; Mismatches 101; Indels 103; Gaps 13;

Qy 9 MSGSXNXXDKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 68  
Db 1 MSGSQNNDKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 60  
Qy 69 RGLALTAATKQAGFASKTETEPVFWYVKEVNLKHELQRFYSLRHHASDVGRGAWLR 128  
Db 61 RGLALTAATKQAGFASKTETEPVFWYVKEVNLKHELQRFYSLRHHASDVGRGAWLR 120  
Qy 129 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLSILFAINIDNK 188  
Db 121 CALNEHSLERYLHMLADRCRLSTFYEDWSFVMDERSMLPTMAAGLSILFAINIDNK 180  
Qy 189 DLNGOSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREITASSAVSILIKPEQETDP 248

Db 181 DLNGOSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREITASSAVSILIKPEQETDP- 239  
Qy 249 LSCPGMSVLMPN-----AKRSGRKRKXPTXSHLMRKMRTLTGCLKRLHGGQRAQR- 301  
Db 240 -----LPVVRNVSADAKCKKXKVTNIIISFDDDEEQNSGDVFKKTPGAGESSED 294  
Qy 302 -----TTPTAPLS-----ISCPLKAPSGLTPME-----SEQOOLMENX 334  
Db 295 NSDRSSVNMISAFEPFGPNSSQSSNSWKIDSLSLNGEFGYQKLDVKSIDDDVDENE 354  
Qy 335 FVPERGVWVPEACEKX-----XRCCKKRRXRVWKLIRKEAQGLGVARE- 382  
Db 355 DDVY-----GNSGRKRGHSESPKPLEGNTC-----LSQMSWAPLKVHLND 398  
Qy 383 -----ATGREHLPLPDAQLSAEGAAQXLRHPLPCQWRGLLPQSRCPKPRGDRDRTRG 436  
Db 399 SDILFPVSGVGSYPADAPLGSLNGTGPEDHVL-----DPGLRYSVEA 443  
Qy 437 PRSPGWSVQCGSOLSRPKSSQPVTSASVPESMTISELRQATVAMNRRKDELEENR 496  
Db 444 -SSPGH-----GSPLS-----SLLPSASVPESMTISELRQATVAMNRRKDELEENR 489  
Qy 497 SLRNLGEMHSAAALROEVDTLKRVABOERQGMKVOALA 538  
Db 490 SLRNLGEMHSAAALROEVDTLKRVABOERQGMKVOALA 531

## RESULT 6

US-09-629-469A-14017  
; Sequence 14017, Application US/09629469A  
; GENERAL INFORMATION:  
; APPLICANT: OTA, TOSHIO  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: NISHIKAWA, TETSUO  
; APPLICANT: HAYASHI, KOJI  
; APPLICANT: SAITO, KAORU  
; APPLICANT: YAMAMOTO, JUNICHI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: OTSUKI, TETSUJI  
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE  
; FILE REFERENCE: 084335/0123  
; CURRENT APPLICATION NUMBER: US/09/629,469A  
; CURRENT FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: JP 1999-248036  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: JP 1999-300253  
; PRIOR FILING DATE: 1999-08-27  
; PRIOR APPLICATION NUMBER: JP 2000-118776  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-183767  
; PRIOR FILING DATE: 2000-05-02  
; PRIOR APPLICATION NUMBER: JP 2000-241899  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/159,590  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 60/183,322  
; PRIOR FILING DATE: 2000-02-17  
; NUMBER OF SEQ ID NOS: 19025  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14017  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-629-469A-14017

Query Match 38.7%; Score 1196; DB 20; Length 375;  
Best Local Similarity 82.2%; Pred. No. 1.4e-95;  
Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;  
Qy 9 MSGSXNXXDKRQFLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 68

Db 1 MSGSQNDKQFLERLLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFAVLQHLKRS 60  
Qy 69 RGLALTAATAIKQAAGFASKTETETEPVFWYVYKEVLNKHLELQRFYSLRHIAADVGGRGAWLR 128  
Db 61 RGLALTAATAIKQAAGFASKTETETEPVFWYVYKEVLNKHLELQRFYSLRHIAADVGGRGAWLR 120  
Qy 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINIDNK 188  
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINIDNK 180  
Qy 189 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSILFREITASSAVSILIKPEQETDPC 248  
Db 181 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSILFREITASSAVSILIKPEQETDPC 239  
Qy 249 LSCPGMSVLPN-----AKRSGRRKRKXPTXSHLMRMKMSRSLTGLTCLKRLHGGRA 299  
Db 240 -----LPVSRNVSADAKCKKRRKKVNTIISFDDDEEQNSGDVFKTPTPGAGES 291

## RESULT 7

US-10-917-503-14017  
; Sequence 14017, Application US/10917503  
; GENERAL INFORMATION:

; APPLICANT: OTA, TOSHIO  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: NISHIKAWA, TETSUO  
; APPLICANT: HAYASHI, KOJI  
; APPLICANT: SAITO, KAORU  
; APPLICANT: YAMAMOTO, JUNICHI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: OTSUKI, TETSUJI  
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE  
; FILE REFERENCE: 084335/0123

; CURRENT APPLICATION NUMBER: US/10/917,503  
; CURRENT FILING DATE: 2004-08-13  
; PRIOR APPLICATION NUMBER: US/09/629,469  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: JP 1999-248036  
; PRIOR FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: JP 1999-300253  
; PRIOR FILING DATE: 1999-08-27  
; PRIOR APPLICATION NUMBER: JP 2000-118776  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: JP 2000-183767  
; PRIOR FILING DATE: 2000-05-02  
; PRIOR APPLICATION NUMBER: JP 2000-241899  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/159,590  
; PRIOR FILING DATE: 1999-10-18  
; PRIOR APPLICATION NUMBER: 60/183,322  
; PRIOR FILING DATE: 2000-02-17  
; NUMBER OF SEQ ID NOS: 19025  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14017  
; LENGTH: 375  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-10-917-503-14017

Query Match 38.7%; Score 1196; DB 35; Length 375;  
Best Local Similarity 82.2%; Pred. No. 1.4e-95;  
Matches 244; Conservative 11; Mismatches 30; Indels 12; Gaps 2;  
Qy 9 MSGSQNDKQFLERLLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFAVLQHLKRS 68  
Db 1 MSGSQNDKQFLERLLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFAVLQHLKRS 60  
Qy 69 RGLALTAATAIKQAAGFASKTETETEPVFWYVYKEVLNKHLELQRFYSLRHIAADVGGRGAWLR 128

Db 61 RGLALTAATAIKQAAGFASKTETETEPVFWYVYKEVLNKHLELQRFYSLRHIAADVGGRGAWLR 120  
Qy 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINIDNK 188  
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINIDNK 180  
Qy 189 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSILFREITASSAVSILIKPEQETDPC 248  
Db 181 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSILFREITASSAVSILIKPEQETDPC 239  
Qy 249 LSCPGMSVLPN-----AKRSGRRKRKXPTXSHLMRMKMSRSLTGLTCLKRLHGGRA 299  
Db 240 -----LPVSRNVSADAKCKKRRKKVNTIISFDDDEEQNSGDVFKTPTPGAGES 291

## RESULT 8

US-10-170-205E-18638  
; Sequence 18638, Application US/10170205E  
; GENERAL INFORMATION:

; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18638  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-18638

Query Match 38.6%; Score 1194; DB 27; Length 250;  
Best Local Similarity 98.7%; Pred. No. 1.2e-95;  
Matches 236; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 9 MSGSQNDKQFLERLLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFAVLQHLKRS 68  
Db 1 MSGSQNDKQFLERLLDAVKQCIQIRFGGRKEIASDSRVTCLCAQFAVLQHLKRS 60  
Qy 69 RGLALTAATAIKQAAGFASKTETETEPVFWYVYKEVLNKHLELQRFYSLRHIAADVGGRGAWLR 128  
Db 61 RGLALTAATAIKQAAGFASKTETETEPVFWYVYKEVLNKHLELQRFYSLRHIAADVGGRGAWLR 120  
Qy 129 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINIDNK 188  
Db 121 CALNEHSLERYLHMLLADRCRLSTFYEDWSFVMDERSSMLPTMAAGLSILFAINIDNK 180  
Qy 189 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSILFREITASSAVSILIKPEQETDPC 247  
Db 181 DLNGQSKFAPTVSDLLKESQNTVTSLLKESQGVSSILFREITASSAVSILIKPEQETDPC 239

## RESULT 9

US-09-758-472-5854  
; Sequence 5854, Application US/09758472  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PH001  
; CURRENT APPLICATION NUMBER: US/09/758,472  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 9632  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5854  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Homo sapiens

```

;
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (198)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (267)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (308)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (312)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-472-5854

Query Match      23.3%; Score 721; DB 22; Length 320;
Best Local Similarity 50.3%; Pred. No. 7.1e-54;
Matches 167; Conservative 21; Mismatches 64; Indels 80; Gaps 7;

QY 110 FYSLRHIASDVGRGAWLRCAALNEHSRLRYLHMLLADRCRLSTFYEDWSFVMDERSMML 169
Db 19 FYSLRHIASDVGRGAWLRCAALNEHSRLRYLHMLLADRCRLSTFYEDWSFVMDERSMML 78

QY 170 PTMAAGLNSILPAINDKDLNGQSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREI 229
Db 79 PTMAAGLNSILPAINDKDLNGQSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREI 138

QY 230 TASSAVSILIKPEQETDPCISCPGMSVLMNP-----AKSGRRKRXKPTXSHLMRMKMS 283
Db 139 TASSAVSILIKPEQETDPCISCPGMSVLMNP-----LPVSVNRVNSADAKCKKRRKKKVTNIISFDDEDE 192

QY 284 RTLGTCLKHLGQGRQARTTPTAPLSISCPPLKAPSLTPMSEQQLMENXFPVFERGVW 343
Db 193 QNSGDXFKKTPGAGESSE-----DNSDRSSVNIMSAFE-SPF 228

QY 344 VPEACEKHXRCXGKRRXRWKLRKEAQGPLG-----VAREA 383
Db 229 GPNS-----NGSQSSNSWKIDSLSLNGEFGYQKLDVKSIDDEDXENEDDDVYGN 278

QY 384 TGREHL-----PLPDAQLSAEG 401
Db 279 SGKRGHSESPRSHWKGTAPFAPDAQVGSXEG 310

RESULT 10
US-10-235-926-5854
; Sequence 5854, Application US/10235926
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PH001CIN
; CURRENT APPLICATION NUMBER: US/10/235,926
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/758,472
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9632
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5854
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:

```

```

; NAME/KEY: misc feature
; LOCATION: (198)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (267)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (308)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-235-926-5854

Query Match      23.3%; Score 721; DB 28; Length 320;
Best Local Similarity 50.3%; Pred. No. 7.1e-54;
Matches 167; Conservative 21; Mismatches 64; Indels 80; Gaps 7;

QY 110 FYSLRHIASDVGRGAWLRCAALNEHSRLRYLHMLLADRCRLSTFYEDWSFVMDERSMML 169
Db 19 FYSLRHIASDVGRGAWLRCAALNEHSRLRYLHMLLADRCRLSTFYEDWSFVMDERSMML 78

QY 170 PTMAAGLNSILPAINDKDLNGQSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREI 229
Db 79 PTMAAGLNSILPAINDKDLNGQSKFAPTVSDLLKSTQNTVTSLLKSTQGVSSLFREI 138

QY 230 TASSAVSILIKPEQETDPCISCPGMSVLMNP-----AKSGRRKRXKPTXSHLMRMKMS 283
Db 139 TASSAVSILIKPEQETDPCISCPGMSVLMNP-----LPVSVNRVNSADAKCKKRRKKKVTNIISFDDEDE 192

QY 284 RTLGTCLKHLGQGRQARTTPTAPLSISCPPLKAPSLTPMSEQQLMENXFPVFERGVW 343
Db 193 QNSGDXFKKTPGAGESSE-----DNSDRSSVNIMSAFE-SPF 228

QY 344 VPEACEKHXRCXGKRRXRWKLRKEAQGPLG-----VAREA 383
Db 229 GPNS-----NGSQSSNSWKIDSLSLNGEFGYQKLDVKSIDDEDXENEDDDVYGN 278

QY 384 TGREHL-----PLPDAQLSAEG 401
Db 279 SGKRGHSESPRSHWKGTAPFAPDAQVGSXEG 310

RESULT 11
US-60-212-356-236
; Sequence 236, Application US/60212356
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE
; FILE REFERENCE: CL000677
; CURRENT APPLICATION NUMBER: US/60/212,356
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 236
; LENGTH: 548
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(548)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-212-356-236

Query Match      22.4%; Score 693.5; DB 37; Length 548;
Best Local Similarity 56.7%; Pred. No. 4.3e-51;
Matches 160; Conservative 22; Mismatches 65; Indels 35; Gaps 5;

```





Mon Jun 13 13:27:10 2005

```
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 30582
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-08631-30582

Query Match      12.5%; Score 388; DB 1; Length 91;
Best Local Similarity 96.3%; Pred. No. 2e-25; 3; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      9 MSGSXNDKRFLLERLLDAVKQCQIRFXGRKEIASDSRVTCLCAQFEAVLQHLKRS 68
Db      1 MSGSQNNDKRFLLERLLDAVKQCQIRFGGRKEIASDSRVTCLCAQFEAVLQHLKRS 60

QY      69 RGLALTAATAIKOAGFASKTET 90
Db      61 RGLALTAATAIKOAGFASKTET 82

Search completed: June 7, 2005, 12:19:32
Job time : 271.467 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 5543.69 Seconds  
(without alignments)  
18065.048 Million cell updates/sec

Title: US-09-155-676B-3  
Perfect score: 2631  
Sequence: 1 cccctcaccagccagcc.....accctgacctcgaggggggg 2631

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

EFT:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_hic:\*
- 4: gb\_est3:\*
- 5: gb\_est4:\*
- 6: gb\_est5:\*
- 7: gb\_est6:\*
- 8: gb\_gsl1:\*
- 9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1770.4	67.3	3971	3	CR749592 Homo sapi
2	938.4	35.7	1640	3	CR612342 full-leng
3	872	33.1	2844	9	AY4113577 Homo sapi
4	711.4	27.0	2765	9	AY4113578 Pan trogl
5	706.2	26.8	911	5	BU195099 AGENCOURT
6	705	26.8	762	4	BI259434 602968032
7	698.6	26.6	4174	3	AK036200 Mus muscu
8	680.6	25.9	1140	4	BM458601 AGENCOURT
9	636.4	24.2	1133	1	AL556130
10	598.2	22.7	2829	9	AY4113579 Mus muscu
11	568.6	21.6	702	7	CR979910 4111556 B
12	568.2	21.6	806	4	BI161002 602865208
13	564.2	21.4	580	5	BP267282 BP267282
14	553	21.0	1095	1	AL578455 AGENCOURT
15	514.2	19.5	955	5	BU541005 AGENCOURT
16	499.8	19.0	909	5	BU541005 AGENCOURT
17	456	17.3	1037	4	CG775354 602649808
18	446.8	17.0	773	7	CK365555 AGENCOURT
19	413.2	15.7	896	5	BU542238 AGENCOURT
20	408.8	15.5	991	4	BM550076 AGENCOURT
21	405.8	15.4	1072	5	BU517750 AGENCOURT
22	391	14.9	701	4	CG774889 602649808
23	389	14.8	768	4	BM048296 603625505
24	386	14.7	774	4	BG822015 602726122

25	385.2	14.6	819	5	BQ231411
26	379.8	14.4	692	4	BI651143
27	368	14.0	702	5	BI651143
28	360.6	13.7	727	4	BI255165
29	348.4	13.2	545	7	N50047
30	343	13.0	615	5	BQ827984
31	341.8	13.0	665	6	CB305346
32	340	12.9	744	6	CA321462
33	335.2	12.7	583	5	BP332147
34	325.6	12.4	1172	4	BM543747 AGENCOURT
35	324.6	12.3	702	7	CN409540 170006001
36	322.2	12.2	579	2	BE407122 601301805
37	321.8	12.2	548	5	BU623799 UI-H-FG1
38	319.8	12.2	702	4	BI489907 603031660
39	316.2	12.0	556	2	BE276256 601144348
40	315.2	12.0	531	6	CA748591 UI-H-FG1
41	314.4	11.9	503	2	BE645094 7e65h06.x
42	312.8	11.9	454	1	AA807939 nu90f09.8
43	312.6	11.9	460	7	W65377 2d32d07.r1
44	311.8	11.9	445	1	AA621781 af06f05.8
45	309.2	11.8	450	4	BI685768 603309421

## ALIGNMENTS

RESULT 1  
CR749592  
LOCUS CR749592 3971 bp mRNA linear HTC 19-AUG-2004  
DEFINITION Homo sapiens mRNA; cDNA DKFZp686J04131 (from clone DKFZp686J04131).  
ACCESSION CR749592  
VERSION CR749592.1 GI:51476844  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3971)  
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Weil,B., Amid,C., Oeinger,A., Fobo,G., Han,M. and Wiemann,S.  
CONSRM The German cDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp686J04131) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering.  
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686J04131  
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.  
FEATURES  
Location/Qualifiers  
1..3971  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686J04131"  
/tissue\_type="uterus"  
/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_sfi; host DH10B; sites SfiIA + SfiIB"  
/dev\_stage="adult"  
/note="MAP3K14 protein, N-terminus truncated, differentially spliced"  
1..3971  
/gene="DKFZp686J04131"  
<1..1302  
/gene="DKFZp686J04131"  
/codon\_start=1  
/product="hypothetical protein"

/protein\_id="CAH18391.1"  
/db\_xref="GI:51476845"  
/translation="CFKQGEGLRALPRSELHLKISPLQLCNHWKLLHHPQDGGPLP  
LPTHPFPFHPWPKEPPLSGFKGLACDSQKPLDPHLSKLACVSDSPKPLFGPHLE  
PSCLSRGAAHEKFSVEYLVAHALQGSVSSGQASHISLTAUKTAAAGSRSPGPKVTEIN  
EGVLITKUKPVDVYREYREVVHATHTQLBLGRGSGEYVHRMEDKQTQCAVAKVKLEL  
FRAEELMACAGTSLRIPLYGAVKAGDPWNIIMELLEGSSGLQLEKQGLCPEDRAL  
VLYQALLEGLEYLHGRIRILYGVADKNVLLSSDGSAAALCDFGHAVCLQDGLGKSL  
TGDIYIPGTTHFAEVLVGRGCDKADVYVSSCMMLHMGHCPTQTFRGPGLCLKVG  
DETQHGCGEACGAGAGRGSCAVADGARGGLVTHVGS"

## ORIGIN

Query Match	67.3%	Score 1770.4	DB 3	Length 3971	
Best Local Similarity	88.8%	Pred. No. 0			
Matches 2291	Conservative	5	Mismatches 195	Indels 88	Gaps 35
QY	1	CCCTCTC	CACAGCCAGGCCATCAAGAGGGGCTGAGAAAGACCCATCCACCGGTGT	60	
Db	1423	CCCTCTCA	CAGCCAGGCCATCAAGAGGGGCTGAGAAAGACCCATCCACCGGTGT	1482	
QY	61	CTCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA	120		
Db	1483	CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA	1542		
QY	121	GCCCTTCGAGGGGAGATATAAAGAACCAAGACATCCACGCCCAATCAAGGCCAATTACC	180		
Db	1543	GCCCTTCGAGGGGAGATATAAAGAACCAAGACATCCACGCCCAATCAAGGCCAATTACC	1602		
QY	181	ACCAGACCCTCATGCCAGCGGAGAGGCTTTCGCAAGGGGCCAGGGCCCGGCCAG	240		
Db	1603	ACCAGACCCTCATGCCAGCGGAGAGGCTTTCGCAAGGGGCCAGGGCCCGGCCAG	1662		
QY	241	CTGAGGAGACAAAGGAGGAGCCCTTAAGCTCCAGGCTCTCTCTCCACACAGAGCCCCCAG	300		
Db	1663	CTGAGGAGACAAAGGAGGAGCCCTTAAGCTCTCAGGCTCTCTCTCCACACAGAGCCCCCAG	1722		
QY	301	AGCCAAAACAAGTCTCCTCCCTTGACTTTGACAAAGGAGGAGTCTGGGATGTGGAAACCT	360		
Db	1723	AGCCAAAACAAGTCTCCTCCCTTGACTTTGACAAAGGAGGAGTCTGGGATGTGGAAACCT	1782		
QY	361	TACCTCTGTCTCCTCTGGAGCAGGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAG	420		
Db	1783	TACCTCTGTCTCCTCTGGAGCAACCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAAG	1842		
QY	421	CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATTATCTCAACAGCCTGT	480		
Db	1843	CAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAAATTATCTCAACAGCCTGT	1902		
QY	481	CCAGCCATTTTCTCTGAGGAGCAGGAGCAATTTCTCTGCTCCTCAGCATCCAGAGCC	540		
Db	1903	CCAGCCATTTTCTCTGAGGAGCAGGAGCAATTTCTCTGCTCCTCAGCATCCAGAGCC	1962		
QY	541	TCTCCTCTGCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAAGCTCGCGGACA	600		
Db	1963	TCTCCTCTGCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAAGCTCGCGGACA	2022		
QY	601	CCCTGAGCTCAGCGGTAACA	CTCTGGAGAGCCAGGCCAGGCTCGAGCTCCAGCTGGA	660	
Db	2023	CCCTGAGCTCAGCGGTAACA	CTCTGGAGAGCCAGGCCAGGCTCGAGCTCGAGCTGGA	2082	
QY	661	ACATGGTGTGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG	720		
Db	2083	ACATGGTGTGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG	2142		
QY	721	TCCAAATACAGTCTCTTAATGGTGAACACCTGCATCCGGGAGTTCACCGGTC	780		
Db	2143	TCCAAATACAGTCTCTTAATGGTGAACACCTGCATCCGGGAGTTCACCGGTC	2202		
QY	781	TGGGAGCATCGCCATCGCATCAGCAGCCAGATCCAGCTTCAGCTTCAGCTTGGTCA	840		
Db	2203	TGGGAGCATCGCCATCGCATCAGCAGCCAGATCCAGCTTCAGCTTCAGCTTGGTCA	2262		
QY	841	CCAAGACGGGAGCCTGTTCCTACGACATGGAGGTGCAGACTCGGGGATTCGACCTTCG	900		

2263	DB	CCAAAGACGGG	CAGCGCTGTTTCGTACGACATGGAGGTCGCAGATCGGGCATTCGACCTGC	2322
901	QY	AGTGCA	CTGGCCCTGATGCGACGCTTCGCTCGAGCTTGAGGGTCAAGCATGGCCAGC	960
2323	DB	AGTGCACATCGG	-CCCTGATGGAGCTTCGCTCGAGCTTGAGGCTCAAGCATGGCCAGC	2381
961	QY	TGAGAACAGGGCCTAACCCCTGCCCTTCCACCGCCGGCTCCACATCGCCGGAAGCAGCT	1020	
2382	DB	TGAGAACAGGGCCTAACCCCTGCCCTTCCACCGCCGGCTCCACATCGCCGG	-AAGCAGGCT	2440
1021	QY	TCCTGCTCGGTGCACATGCTGCCCTTGAGAAACACAGGCTCAGCCGTTCCAGGGGATVTG	1080	
2441	DB	TCCTGCTCGGTGCACATGCTGCCCTG	-AAAACAGGCTCAGCCGTTCCAGGG--ATC	2496
1081	QY	NCCAGCCCCCGGCTCARCAGNTGGGAACAGGGGCTCGNACGNAGCNAAAGTTNGGGGG	1140	
2497	DB	TGCAGGCCCGGCTCAGCAGTGGGACAGGGCCTCGCAGCAGCAAG	-----GTGGGG	2549
1141	QY	CAAGCNAGAAATGCTCCAGGATTTCA	NCCTTGAGCCCTGCTGCTGAADA	1200
2550	DB	GCAAGCAGAAATGCTCCAGGATTTCA	ACC---TGAGCCCTGCCCCACCTCTGCTAGAA	2606
1201	QY	AAACAYTNCGCCACGCTGAAGACAGACAGAGGAGTAGCNACGAGTGTNNACCTYGGGAA	1260	
2607	DB	AACACT--CCGCCACGTGAAGACA	----GAGGAGTAGGCAGGATTTACCTCGGAAA	2660
1261	QY	ACAAAACAGGGATCTTTNTTCTGCCCTGCTCCAGTNCAGTTCGAGTTGGCTGNA	CCGCTTGG	1320
2661	DB	CAAAACGATCTTCTGCTGCCCTGCTCAG	-----TCGAGTTGGCTGACCCGCTT--G	2711
1321	QY	ANTCAGTGA	CCATTTGTTGTCACANACAGGGGAGCAGCTTCACGCTCGGTGAGAAAGG	1380
2712	DB	GATCAGTGA	CCATTTGCTGTCAGACAGGGAGCAGCTTCACGCTCGGTGAGAAAGG	2770
1381	QY	GTGGCGCAGCCCTTCGGCCCTCACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA	1440	
2771	DB	GTGGCGCAGCCCTTCGGCCCTCACCT	-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA	2828
1441	QY	AGGNCCAAANCTCAGGNTTCAGTCAGAACACAGTNCAGCAGGTATGCCGCCCGNTA	1500	
2829	DB	AGGGCCCAAACTC	---AGGTTCAGTGCAGAACACAGGT-CHAGAGGTATGCCGCCCGTAG	2884
1501	QY	GTTTAA	NNGGGGCCCTCTNAAACCCCTTGCTTNGGCTNACCTTNGGCGAGCTCANCC	1560
2885	DB	GTATA	----GGGGCCCTCTAAACCCCTTGCTTGGCTCA--CCTGGCCAGCTCA-CCC	2936
1561	QY	CTTTGGGTG	TAGGGGAAAAGATGCTGACCTCGGAGGCTTCCTGTGTAAGATACAC	1620
2937	DB	CTTTGGGTG	TAGGGGAAAAGATGCTGACCTCGGAGGCTTCCTGTGTAAGATACAC	2995
1621	QY	CACACTTTTCAGTGTGTGAAACACAGTCTCAGTGTGACCTCTGTTTCAGCCAAAGGAC	1680	
2996	DB	CACACTTTTCAGTGTGTGAAACACAGTCTCAGTGTGACCTCTGTTTCAGCCAAAGGAC	3055	
1681	QY	AAAGAAGTGTGTAAGTGAAGTGTCTCAGTNC	CCAGACATGTGCCCTTTGCTGCTG	1740
3056	DB	AAAGAAGTGTGTAAGTGAAGTGTCTCAGT	-CCCCAGACATGTGCCCTTTGCTGCTG	3114
1741	QY	GCTACCACTCTTCCCGCAGACAGACGCCCGAGCCCTTCAGGCCCAGACATGCCCCCAG	1800	
3115	DB	GCTACCACTCTTCCCGCAGACAGACGCCCGAGCCCTTCAGGCCCAGACATGCCCCCAG	3174	
1801	QY	ACTCGCTGGCACTCAGTTCCTCATCTGTATAAGTGAAGGTGATG	CAGGATATGCTGTA	1860
3175	DB	ACTCGCTGGCACTCAGTTCCTCATCTGTATAAGTGAAGGTGATG	CAGGATATGCTGTA	3234
1861	QY	CAGGAA	CAGTCTGTGGATGGACATGATCAGTGTCTNAGNAAACACAGACAGACGACG	1920
3235	DB	CAGGAA	CAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAACAGACAGACAGACG	3292
1921	QY	TCGGCGCCCCAGNCCCCACTNATCAGTGTNCCACGCTGTNGGTTNCCCAGNAGACA	1980	

3293	TCCGGGCCCCAGCCCCCACTATCAGTGT---CCAGCGGTGCTGGTTCC-----CAGA	3342
1981	GCTNCAGNCATCANCACTGACACTNCACCCCTNGCCCTGCCCTTGGCCAGGAGGCTACTG	2040
3342	GCACAGCTCAGCATCACTGACACTCACCTGCCCTGCCCTGG--CCAGAGGCTACTG	3399
2041	CCGNA CGGCACTTTTGACATCTGATGACCTCAAAGCACTTTTCATGGCTTNGCCCTCTNN	2100
3400	CCG-ACGGCACTTTGGCACT--CTGATGACCTCAAAGCACTTTTCATGGCTCTCTGGCA	3456
2101	GCAGGNCAGGNCAGGNCAGTGCANCTGTAGGAGCATANGCAANGCCAGGATGG	2160
3457	GGGAGGGCAGGGCAGTGCACCT-----GTAGGAGCATAGCAAGCCAGGATGG	3506
2161	GGTNGAAGGANCACAGCTCTTGAGCTGTCCANCACTGATGTGACTTNCCTCAAACTCTT	2220
3507	GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT--CCTCAAACCTCTT	3561
2221	NCCAGNATTTCTTAAGAATPAGCANGCCCCCTTNCCTCATATGTCCTTGGCTCTTCT	2280
3562	--CCAGATTTCTCTAAGAAATAGCACCCCC--TTCCCCATTTGCCCACTTAGCTCTTCT	3617
2281	CCCAGGGGAGCTANCTCAGGACTCAGTACATTAATCAGCTGTGNAATCGTCAGGGGG	2340
3618	CCCAGGGGAGCTA-CTCAGGACTCAGTACATTAATCAGCTGTG-AAATCGTCAGGGGG	3675
2341	TGCTCTGACCTCAACCTCCTGGGGCAGGGGACGGCAGACTCCGTGGGAGAAAGCTCAT	2400
3676	TGCTCTGACCTCAACCTCCTGGGGCAGGGGACGGCAGACTCCGTGGGAGAAAGCTCAT	3735
2401	TCCCACATCTTGGCAAGACAGCTTTTNGTCCAGCTGTCCATATGATGACAGCTGCTCCC	2460
3736	TCCCACATCTTGGCAAGACAGCTTTT-GTCCAGCTGTCCCATTTGATGACAGCTGCTCCC	3794
2461	GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCTTGGTACTGACAGCTCTG	2520
3795	GGGGAGAGAGCCCGGGCCCCCAGCACATAAAGAACTGCAGCTTGGTACTGACAGCTCTG	3854
2521	GGTTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGATCACAATAATGTTAAAAA	2579
3855	GGTTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGATCACAATAATGTTAAAAA	3913
RESULT 2		
CR612342	1640 bp mRNA linear HTC 21-JUL-2004	
LOCUS	full-length cDNA clone CS0DK011W14 of Hela cells Cot 25-normalized	
DEFINITION	of Homo sapiens (human).	
ACCESSION	CR612342.1 GI:50493149	
VERSION	HTC; CDSLT cDNA.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 1640)	
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE	Full-length cDNA libraries and normalization	
JOURNAL	Unpublished	
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue	
REFERENCE	2 (bases 1 to 1640)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)	
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.	

```
QY 1724 GTGCCCTTTGCTGCTGCTTACCACTCTTCCCGACAGCAGGCCCCGAGCCCTTTCAG 1783
Db 854 GTGCCCTTTGCTGCTGCTTACCACTCTTCCCGACAGCAGGCCCCGAGCCCTTTCAG 913
QY 1784 GCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTCTGTAAGGTGAAGGGTG 1843
Db 914 GCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTCTGTAAGGTGAAGGGTG 973
QY 1844 ATGCAGGATATGCTGTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTTAAAGNAAA 1903
Db 974 ATGCAGGATATGCTGTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTTAAAGNAAA 1031
QY 1904 GCAGCAGAGAGAGCTGTCGGGGCCCCAGNCCCCACTNATCAGTGTNCCAGCGTGTG 1963
Db 1032 GCAGCAGAGAGAGCTGTCGGGGCCCCAGNCCCCACTNATCAGTGTCTTAAAGNAAA 1088
QY 1964 GTTNCCTCCAGNAGCAGCTNCACTNCACTNCACTNCACTNCACTNCACTNCACTNCACTN 2023
Db 1089 TTCCCC-----CAGAGCACAGCTCAGCATCACACTGACCTCACCCTGCCCCCT 1140
QY 2024 NGGCCANGAGGTACTGCGCAGGACCTTGGACATCTGATGATGATGATGATGATGATGATGAT 2083
Db 1141 GG--CCAGAGGCTACTGCGG--ACGGCACTTTGCAC--TCTGATG--ACCTCAAAGCACTTTC 1195
QY 2084 ATGGCTNGCCTCTNNGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 2143
Db 1196 ATGGCTNGCCTCTGCGCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 1245
QY 2144 GCAAGCCAGGAGATGGGCTGNAAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 2203
Db 1246 AGCAAGCCAGGAGATGGGCTGNAAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG 1302
QY 2204 CTNCCCTCAAACCTCTTNNCCAGNATTTCTTAAGATAGCAGGNCAGGNCAGGNCAGGNCAGG 2263
Db 1303 CT--CTCAAACCTCTT--CCAGATTTCTTAAGATAGCAGGNCAGGNCAGGNCAGGNCAGG 1356
QY 2264 CAGCTTAGCTCTTCTCCAGGAGGAGTANCTCAGGACTCAGGTAGCATTAATCAGCT 2323
Db 1357 CAGCTTAGCTCTTCTCCAGGAGGAGTANCTCAGGACTCAGGTAGCATTAATCAGCT 1415
QY 2324 GTGNAATCGTCAGGAGGCTGTCTGCTAGCTCTCACTCTGCGGAGGAGGAGGAGGAGGAGG 2383
Db 1416 GTG--AATCGTCAGGAGGCTGTCTGCTAGCTCTCACTCTGCGGAGGAGGAGGAGGAGGAG 1474
QY 2384 CCGTGGAGAGCTCATTCACATCTTGCACAGACAGGCTTTGTCAGCTGTCCACAT 2443
Db 1475 CCGTGGAGAGCTCATTCACATCTTGCACAGACAGGCTTTGTCAGCTGTCCACAT 1533
QY 2444 TGAGTCAGACTGCTCCCGGGAGAGAGGCCCCGGCCCCCAGCAGCATAAAGACTGACGCT 2503
Db 1534 TGAGTCAGACTGCTCCCGGGAGAGAGGCCCCGGCCCCCAGCAGCATAAAGACTGACGCT 1593
QY 2504 TGCTACTGACAGCTCTGGTGTGTAGAGAACTCTTTGTAAGCAATAAA 2550
Db 1594 TGCTACTGACAGCTCTGGTGTGTAGAGAACTCTTTGTAAGCAATAAA 1640

RESULT 3
AY413577 2844 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413577
VERSION AY413577.1 GI:39769539
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2844)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D., and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2844)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D., and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..2844
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2844
/gene="MAP3K14"
/locus_tag="HCM4937"

ORIGIN
Query Match 33.1%; Score 872; DB 9; Length 2844;
Best Local Similarity 100.0%; Pred. No. 3.8e-216;
Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 TGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAA 165
Db 1973 TGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAA 2032
QY 166 ATCAGCCAAATTACACAGACCCCTCCATGCCAGCCGAGAGAGCTTTGCGAAGGGGCC 225
Db 2033 ATCAGCCAAATTACACAGACCCCTCCATGCCAGCCGAGAGAGCTTTGCGAAGGGGCC 2092
QY 226 CAGGCCCCGGCAGCTGAGGAGACAAAGGAGAGCCCTTAAGCTCAGAGCTCTCTTCC 285
Db 2093 CAGGCCCCGGCAGCTGAGGAGACAAAGGAGAGCCCTTAAGCTCAGAGCTCTCTTCC 2152
QY 286 CACAGAGCCCCAGAGCCAAACAAAGTCTCTCCCTTGAATTTGAGCAGGAGGAGTCTG 345
Db 2153 CACAGAGCCCCAGAGCCAAACAAAGTCTCTCCCTTGAATTTGAGCAGGAGGAGTCTG 2212
QY 346 GGATCTGGAGACCCCTTACCTCTCTCTGAGCCAGCCCTCCAGAAACCCAGCT 405
Db 2213 GGATCTGGAGACCCCTTACCTCTCTCTGAGCCAGCCCTCCAGAAACCCAGCT 2272
QY 406 CACAGAGCGGAAACAAACCGTCCCGAGCAGGAACTGCAGAGCTGGAATATAT 465
Db 2273 CACAGAGCGGAAACAAACCGTCCCGAGCAGGAACTGCAGAGCTGGAATATAT 2332
QY 466 TCTCAAAGCTGTCCAGCCATTTCTTGGAGAGCAGGAGCAATTTCTCTGTCGC 525
Db 2333 TCTCAAAGCTGTCCAGCCATTTCTTGGAGAGCAGGAGCAATTTCTCTGTCGC 2392
QY 526 TCAGCATCGACAGCTCTCCCTGTCGATGACAGTGAAGAAACCCATCAAAGGCTCTC 585
Db 2393 TCAGCATCGACAGCTCTCCCTGTCGATGACAGTGAAGAAACCCATCAAAGGCTCTC 2452
QY 586 AAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTTGGAGAGCAGCAGGCGGAGT 645
Db 2453 AAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTTGGAGAGCAGCAGGCGGAGT 2512
QY 646 GAAGCTCGAGCTGGAAACATGCTGTCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 705
Db 2513 GAAGCTCGAGCTGGAAACATGCTGTCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2572
QY 706 TCAATGGTGTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCCGGAGT 765
Db 2573 TCAATGGTGTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCCGGAGT 2632
```

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D., and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 2844)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D., and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.

## FEATURES

Location/Qualifiers

source

1..2844

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

&lt;1..&gt;2844

/gene="MAP3K14"

/locus\_tag="HCM4937"

## ORIGIN

Query Match 33.1%; Score 872; DB 9; Length 2844;

Best Local Similarity 100.0%; Pred. No. 3.8e-216;

Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 TCACCGGGTCAAAGTGGAGAGATCGCCACTGGGATCAGAGCCAGATCCAGCTGCAG 825  
 |||||  
 Db 2633 TCACCGGGTCAAAGTGGAGAGATCGCCACTGGGATCAGAGCCAGATCCAGCTGCAG 2692  
 |||||  
 QY 826 CTTTCAGCTTGGTCAACAAAGAGCGGAGCGCTGTTTCGCTAGACATGAGGTGCCAGACT 885  
 |||||  
 Db 2693 CTTTCAGCTTGGTCAACAAAGAGCGGAGCGCTGTTTCGCTAGACATGAGGTGCCAGACT 2752  
 |||||  
 QY 886 CGGCATCGACCTGCAGTGCACACTGGCCCTGTGATGGCAGCTTCGCTGGAGCTGGAGGG 945  
 |||||  
 Db 2753 CGGCATCGACCTGCAGTGCACACTGGCCCTGTGATGGCAGCTTCGCTGGAGCTGGAGGG 2812  
 |||||  
 QY 946 TCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA 977  
 |||||  
 Db 2813 TCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA 2844  
 |||||

RESULT 4  
 AY413578 2765 bp DNA linear GSS 17-DEC-2003  
 LOCUS Pan troglodytes MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 ACCESSION AY413578  
 VERSION AY413578.1 GI:39769540  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 1 (bases 1 to 2765)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1..2765  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 <1..>2765  
 /gene="MAP3K14"  
 /locus\_tag="HCW4937"

Query Match 27.0%; Score 711.4; DB 9; Length 2765;  
 Best Local Similarity 82.0%; Pred. No. 3.7e-174;  
 Matches 715; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 106 TGGGAGGTCTAGAGCCCTTGGAGGGGAGAGATATAAGAACCAAGACATCCCGCCAA 165  
 |||||  
 Db 1894 TGGGAGGTCTAGAGCCCTTGGAGGGGAGAGATATAAGAACCAAGACATCCCGCCAA 1953  
 |||||  
 QY 166 ATCAAGCAATTACCAACAGACCTTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCC 225  
 |||||  
 Db 1954 ATCAAGCAATTATCACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGGCC 2013  
 |||||  
 QY 226 CAGGCCCGGCCAGCTGAGGAGACAAACAGCGAGGCCCTTAAGCTCAGCTCCTCTCC 285  
 |||||

Db 2014 CAGGCCCGGCCAGCTGAGGAGACAACAGCAGAGAGCCCTTAAGCTCAGAGCTCCTCTCC 2073  
 |||||  
 QY 286 CACCAGAGCCCCAGAGCCAAACAAGTCTCTCCCTTGAATTTTGAGCAAGAGAGTCTG 345  
 |||||  
 Db 2074 CACCANAGCCCCAGAGCCAAACAAGTNTCTCTCCCTTGAATTTTGAGCAAGAGAGTCTG 2133  
 |||||  
 QY 346 GGATGTGGAAACCTTACCT 405  
 |||||  
 Db 2134 GGATGTGGAAACCTTACCT 2193  
 |||||  
 QY 406 CACCAGAGCGGAAAGAACCGTCCCGAGCAGGAACCTGACAGCTGCAAAATAGAAATAT 465  
 |||||  
 Db 2194 CACCAGAGCGGAAAGAACCTTCCCGAGCAGGAACCTGACAGCTGCAAAATAGAAATAT 2253  
 |||||  
 QY 466 TCCTCAACAGCCTGTCCAGCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 525  
 |||||  
 Db 2254 TCCTCAACAGCCTGTCCAGCCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2313  
 |||||  
 QY 526 TCAGCATCGACAGCCT 585  
 |||||  
 Db 2314 TCAGCATCGACAGCCT 2373  
 |||||  
 QY 586 AAAGCTCGCGGACACCTCTGAGCTCAGCGTACACTCTCTCTCTCTCTCTCTCTCTCTCT 645  
 |||||  
 Db 2374 NNN 2433  
 |||||  
 QY 646 GAAGCTCCAGCTGGAACATGCTGTGCGCCCGGGGCGGCCACCCAGCTATT 705  
 |||||  
 Db 2434 NNN 2493  
 |||||  
 QY 706 TCAATGGTGTGAAAGTCAAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGAGT 765  
 |||||  
 Db 2494 NNNNNNGTGTGAAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCGGAGT 2553  
 |||||  
 QY 766 TCACCGGGTCAAGTGGAGACATCCGCTGTCATCAGCAGCAGCAGATCCAGCTGCAG 825  
 |||||  
 Db 2554 TCACCGGGTCAAGTGGAGACATGCGCACTGGCAGCAGCAGCAGATCCAGCTGCAG 2613  
 |||||  
 QY 826 CTTTCAGCTTGGTCAACAAAGAGCGGAGCGCTGTTTCGCTACGATGAGGTGCCAGACT 885  
 |||||  
 Db 2614 CTTTCAGCTTGGTCAACAAAGAGCGGAGCGCTGTTTCGCTACGATGAGGTGCCAGACT 2673  
 |||||  
 QY 886 CGGCATCGACCTGCAGTGCACACTGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGG 945  
 |||||  
 Db 2674 CGGCATCGACCTGCAGTGCACACTGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGG 2733  
 |||||  
 QY 946 TCAGCATGGCCAGCTGGAGAACAGGCCCTTAA 977  
 |||||  
 Db 2734 TCAAGCATGGCCAGCTGGAGAACAGGCCCTTAA 2765  
 |||||

RESULT 5  
 BUI95099 911 bp mRNA linear EST 04-SEP-2002  
 LOCUS AGENCOURT\_7964504 NIH\_MGC\_72 Homo sapiens cdna clone IMAGE:6169272  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BUI95099  
 VERSION BUI95099.1 GI:22709083  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 911)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: ATCC/DCTD/DTP  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM1534 row: c column: 01  
High quality sequence stop: 556.  
Location/Qualifiers

## FEATURES

source

1..911  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6169272"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

## ORIGIN

Query Match 26.8%; Score 706.2; DB 5; Length 911;  
Best Local Similarity 91.3%; Pred. No. 5.9e-173;  
Matches 837; Conservative 5; Mismatches 58; Indels 17; Gaps 9;

QY 420 GCAACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAATTATTCCTCAACAGCCTG 479  
DB 1 GCACCGTCCCGAGCAGGAACTGCAGCAGCTGGAATAGAATTATTCCTCAACAGCCTG 60  
QY 480 TCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATCTCTCTGCTCTCAAGATCGACAGC 539  
DB 61 TCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATCTCTCTGCTCTCAAGATCGACAGC 120  
QY 540 CTCTCCCTCGGATGACAGTGCAGAGACCCATCAAGGCCCTCAAGGCTCGGGGAC 599  
DB 121 CTCTCCCTCGGATGACAGTGCAGAGACCCATCAAGGCCCTCAAGGCTCGGGGAC 180  
QY 600 ACCTGAGCTCAGGCTGACACTCTCTGGAGCAGCCAGGCTCGAAGCTCCAGCTGG 659  
DB 181 ACCTGAGCTCAGGCTGACACTCTCTGGAGCAGCCAGGCTCGAAGCTCCAGCTGG 240  
QY 660 AACATGGTGTGCGCCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGTTGAA 719  
DB 241 AACATGGTGTGCGCCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGTTGAA 300  
QY 720 GTCCAAATACAGTCTCTTAATGTGAACACCTGCACATCGGAGTTCCACCGGGTCAA 779  
DB 301 GTCCAAATACAGTCTCTTAATGTGAACACCTGCACATCGGAGTTCCACCGGGTCAA 360  
QY 780 GTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTC 839  
DB 361 GTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTC 420  
QY 840 ACCAAAGAGCGGAGCCTGTTGCTAGCAGATGGAGTGCAGACTCGGGGATCGACCTG 899  
DB 421 ACCAAAGAGCGGAGCCTGTTGCTAGCAGATGGAGTGCAGACTCGGGGATCGACCTG 480  
QY 900 CAGTGACACATGGCCCTGATGTCAGCTTCGCTGGAGCTGGAGGTCAGAGATGCCAC 959  
DB 481 CAGTGACACATGGCCCTGATGTCAGCTTCGCTGGAGCTGGAGGTCAGAGATGCCAC 540  
QY 960 CTGGAGAACAGGCCCTTAACCCCTGCCACCGCGGCTCCACACTCGCGGAAAGCAGCC 1019  
DB 541 CTGGAGAACAGGCCCTTAACCCCTGCCACCGCGGCTCCACACTCGCGGAAAGCAGCC 599  
QY 1020 TTCTCTGCTGGTGCAGATGCTGCTGAAACACAGGCTCAGCGCTTCCTCCAGGGGATYT 1079  
DB 600 TTCTCTGCTGGTGCAGATGCTGCTGAAACACAGGCTCAGCGCTTCCTCCAGGGGATCT 658  
QY 1080 GNCACGCCCCGGCTCARGAGNTGGAAACAGGGGCTTCGNCAGCAGCAGGAGGTTGGG 1139  
DB 659 GGCAG - CCCCCGGCTCAGAGTGGGACAGGGGCTCGGAGCAGC-----AAGTGGGG 712  
QY 1140 GCAAGCAGAGATGCCTCCAGGATTTTCAACANCTGAGCCCTGCCCCCAGCCCTGCTGA 1199

DB 713 GCAAGCAGAGATGCCTCCCGAGATTTACACCTTGAGCCTT-GGCCACCCCTGGCTGAG 770  
QY 1200 AAAACAYTNCGCCACAGTGAAGACAGAGAGATGCGAGGATGTTNACCTYGGGA 1259  
DB 771 AAAACACTCCGCCACAGTGAAGACAGAGATGCGAGGATGTTNACCTYGGGA 826  
QY 1260 AACAAACAGGATCTTTTCTGCCCCTGCTCCAGTTCGAGTT-GGCTTGNACCCGCTT 1318  
DB 827 AAAACACAGGATCTTTTCTGCCCCTGCTCCAGTTCGAGTTGGGCTGACCCGCTT 884  
QY 1319 GGANTCAGTACCAATTT 1335  
DB 885 TGGATCCGTGACCAATT 901  
RESULT 6  
BI259434 762 bp mRNA linear EST 17-JUL-2001  
LOCUS 602968032F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:5107807 5',  
DEFINITION mRNA sequence.  
ACCESSION BI259434  
VERSION BI259434.1 GI:14816769  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 762)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11260 row: o column: 08  
High quality sequence stop: 760.  
Location/Qualifiers  
1..762  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5107807"  
/tissue\_type="cervical carcinoma cell line"  
/lab\_host="DH10B"  
/clone\_lib="NIH MGC 12"  
/note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN

Query Match 26.8%; Score 705; DB 4; Length 762;  
Best Local Similarity 97.6%; Pred. No. 1.1e-172;  
Matches 737; Conservative 0; Mismatches 15; Indels 3; Gaps 2;

QY 319 CCTTGACTTTGAGCAAGGAGGAGTCTGGATCTGGGAACCTTACCTCTGTCTCCTCG 378  
DB 7 CGTGGCTTTGAGCAAGGAGGAGTCTGGATCTGGGAACCTTACCTCTGTCTCCTCG 66  
QY 379 AGCCAGCCCTTCCGAGAAACCCAGCTACACAGCGGAAAGCAACCTGCCGAGCAG 438  
DB 67 AGCCAGCCCTTCCGAGAAACCCAGCTACACAGCGGAAAGCAACCTGCCGAGCAG 126  
QY 439 AACTCAGCAGCTGGAATAGAAATTATTCCTCAACAGCCTGCCAGCCATTTTCTCTG 498  
DB 127 AACTGCAAGCAGCTGGAATAGAAATTATTCCTCAACAGCCTGCCAGCCATTTTCTCTG 186





PYMLDSAVLDKLAGVSGQRLPPPHLSQLAHGDQKQPLPGPHLESSCPSSRGALKEVP  
VEYLIVHALQWSSVSSQAHSLAKTWSGSAKQRLQGPETEDNEGVLTEKLPVD  
YEVRELVHMTQPRVGRSGFGEVHRMDKQTFQCAVKKVLLEVPRVEVLVACAGLS  
SRIVPLICAVREGFWNIFMELLEGSLGQIKOMGCLIPEDRALYILGQALEGLEYL  
HTRILHGDVADKADNLLSSDGRALCDFGHALCLOPDLGKSLITGDIYIPGETHMA  
PEVWKGPCDAKVDIWSGCCMLHMLNGCHPWTFYRGPCLCIKIASPPPIREIPSPC  
APLTAQAIQEGIRKBPVHRASAMELRKVKALQVGLKSPWKGEYKPRPPQDA  
TCHQTLPTPRENPAPAKANTDGAPEPPPLPEPPSKAPALNLSKEESGTWELPL  
SLSDPATKGPSFPDRRATLPELQQLLELFLNLSLQPSLEQELILSLDLSL  
SLSDSEKNPSKASOSSRDLSLSSGVHNSWNSQAEARTCSCTALARGRPDIPIFYNGV  
KVQIQSLGHEHJHIREFHRVKVDTATGISSQIPATAFSLVTKDQOPCYDMEVPSG  
IDLQCTLPADPGSFATWVRKHQLENRP"  
4157. .4162  
/note="putative"  
polyA\_signal  
4174  
/note="putative"

## ORIGIN

Query Match 26.6%; Score 698.6; DB 3; Length 4174;  
Best Local Similarity 80.7%; Pred. No. 9.3e-171;  
Matches 850; Conservative 0; Mismatches 179; Indels 24; Gaps 2;  
QY 1 CCCCTCTCAGCCAGCCATCCAAAGAGGGGTGAGGAAAGAGCCCATCCACCGCGT 60  
DB 1958 CACCCCTCAGCCAGCCAGCCATCCAAAGAGGGGTGAGGAAAGAGCCCGTCCACCGAGCAT 2017  
QY 61 CTGACGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
DB 2018 CTGCCATGGAGCTTGGAGGAAGTGGGCAAGGCATCAGGAAGTGGGAGGTCTGAAGA 2077  
QY 121 GCCCTTGGAGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGGCCAATTACC 180  
DB 2078 GCCCTTGGAGGAAGGAATATAAGAACCAAGACCTCCACCCCAAGCAAGCCACCTGCC 2137  
QY 181 ACAGACCTCTCATGCCAGCGAGAGAGCTTTGGCAAGGGCCCGAGGCGCCCGGCGCAG 240  
DB 2138 ACCAGACCTTACTACTCCGCGCGAGAGA-----GAAACCCACCCAG 2176  
QY 241 CTGAGGAGACACAGCAGAGCCCTTAAGCTCCAGCTCCTCTCCACACAGAGCCCGCCAG 300  
DB 2177 CCAAGGCCAACAGAGCGGGGTCTTGAGCCTCAGCCTCCTCTACGCCAGAACACCCAG 2236  
QY 301 AGCCAAACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCT 360  
DB 2237 AACCGAGCAAGCGCAGCCCTGAACCTGAGCAAGGAGAGTCTGGCACATGGGAACCC 2296  
QY 361 TACCTCTGTCTCCCTGGAGCAGCCCTGCGCAGAAACCCAGCTCACAGAGCGGAAAG 420  
DB 2297 TGCCTCTGTCTCCCTGGAGCCAGCCACTGCCAAAGGCCCGCAGCTTCCAGACCGGAGG 2356  
QY 421 CAACCGTCCCGAGCAGGAACCTGACAGCTCGAAATAGAAATTATCTCAACAGCCGTGT 480  
DB 2357 CAACCTTGCAGAGCTGAGCTTACAGCAACTGGAGATAGAACTGTTCTCAACAGCCGTGT 2416  
QY 481 CCGAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTCCCTCAGCATTCGACAGCC 540  
DB 2417 CCGAGCCGTCTCTCTGGAGGAACAGGAACAAATCTCTCTCGCTCAGCATTCGACAGCC 2476  
QY 541 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGGACA 600  
DB 2477 TCTCCGTGTGATGACAGTGAAGAAATCCATCGAAGGCCCTCTCAGAGCTCACCGGACA 2536  
QY 601 CCTGAGCTCAGCGGTACACTCTCGAGCAGCAGCGGAGCTCGAAGCTCCAGCTCGA 660  
DB 2537 CCTGAGTCTGCGGTGCACTCTTGGACAGCAGCAAGCTGAGCAAGAACCTGCACTGCA 2596  
QY 661 ACATGGTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 720  
DB 2597 GCACGGCGTGGCGGGGGCGGCTACTTGACATCCCGAGCTACTTCAACGGGGTCAAGG 2656  
QY 721 TCCAAATCAGTCTCTTAATGTGAACACCTGCAATCCGGGAGTTCACCGGGTCAAG 780  
DB 2657 TCCAGATCCAGTCTCTCAATGGCGAACACCTTCATATCCGGGAATTTCACCGCGTCAAGG 2716

QY 761 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTTCAGCTTCAGCTTGGTCA 840  
DB 2717 TGGGAGACATTCGCCACCGGCATCAGCAGCCAGATCCAGCCACAGCTTCAGCTTGGTGA 2776  
QY 841 CCAAAGACGGGAGCGCTCTTGGCTACGACATGGAGGTGCCAGACTCGGGGCATTCGACCTGC 900  
DB 2777 CCAAAGATGGACAGCGCTCTTGGCTATGACATGGAGGTGCCAGACTCGGGGCATTCGACCTGC 2836  
QY 901 AGTGACACATGGCCCCCTGATGGAGCTTCGCTGGAGGTGAGGAGTCAAGCATGGCCAGC 960  
DB 2837 AGTGACACCTGGCCCCCTGATGGAGCTTCGCTGGAGGTGAGGAGTCAAGCATGGCTCAGC 2896  
QY 961 TGGGAACAGAGCCCTTAACCTGCCCTCCACCGCGGTCCACACTG---CCGGAAGCAG 1017  
DB 2897 TGGGAACAGAGCCCTTAAGCTCTGTCTTCCACACACAGCTCCACTCTGGCCACGGCAGGCTT 2956  
QY 1018 CTTCTCTCTCGGTGCACGATGCTGCCCTGAAA 1050  
DB 2957 CTTTGTATCTCAGTGTCTGCTGGTGTCTGAGA 2989  
RESULT 8  
BM458601  
LOCUS  
DEFINITION BM458601 1140 bp mRNA linear EST 05-FEB-2002  
AGENCOURT 6413524 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:5497658  
5' mRNA sequence.  
ACCESSION BM458601  
VERSION BM458601.1 GI:18507641  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1140)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM12128 row: c column: 03  
High quality sequence stop: 724.  
FEATURES  
Location/Qualifiers  
source 1..1140  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:5497658"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_85"  
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.867 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 25.9%; Score 680.6; DB 4; Length 1140;  
Best Local Similarity 83.2%; Pred. No. 3.1e-165;  
Matches 938; Conservative 4; Mismatches 144; Indels 41; Gaps 15;  
QY 458 AGAATATTCTCTCAACAGCCGTGCCCGCAATTTCTCTGAGGAGGAGGAGCAATCT 517  
DB 44 AGAATATTCTCTCAACAGCCGTGCCCGCAATTTCTCTGAGGAGGAGGAGCAATCT 103

```
Qy 518 CTCGTGCTCAGCATCGACAGGCTCTCCCTGTGCGATGACAGTGAAGAACCCTCAAA 577
Db 104 CTCGTGCTCAGCATCGACAGGCTCTCCCTGTGCGATGACAGTGAAGAACCCTCAAA 163
Qy 578 GGCCTCTCAAAAGCTCGCGGACACCTCAGCTCAGCGGTACACTCTCGAGCAGCCAGGC 637
Db 164 GGCCTCTCAAAAGCTCGCGGACACCTCAGCTCAGCGGTACACTCTCGAGCAGCCAGGC 223
Qy 638 CGAGGCTCGAAGTCCAGCTCGAACAATGCTGTGCCCCGGGGCGGCCACCGACACCCC 697
Db 224 CGAGGCTCGAAGTCCAGCTCGAACAATGCTGTGCCCCGGGGCGGCCACCGACACCCC 283
Qy 698 AAGCTATTTCATGCTGAAGTCCAAATACAGTCTCTTAATGGTGAACACTGCGCAT 757
Db 284 AAGCTATTTCATGCTGAAGTCCAAATACAGTCTCTTAATGGTGAACACTGCGCAT 343
Qy 758 CCGGAGGTTCCACCGGGTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCC 817
Db 344 CCGGAGGTTCCACCGGGTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCC 403
Qy 818 AGCTGACGCTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGACATGGAGGT 877
Db 404 AGCTGACGCTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGACATGGAGGT 463
Qy 878 GCGAGACTCGGGCATCGACCTGCGAGTGCACACTGGCCCCCTGATGCGACTTGGCTGGAG 937
Db 464 GCGAGACTCGGGCATCGACCTGCGAGTGCACACTGGCCCCCTGATGCGACTTGGCTGGAG 523
Qy 938 CTGAGGGGTCAAGCATGCGCAGCTGGAGAACAGAGCCCTTAACCTGCCCTCCACCGCGGC 997
Db 524 CTGAGGGGTCAAGCATGCGCAGCTGGAGAACAGAGCCCTTAACCTGCCCTCCACCGCGGC 583
Qy 998 TCCACACTGCGGAAAGCAGGCTTCTCTGCTCGGTGACAGTGTGCTGCTGAAACACAGG 1057
Db 584 TCCACACTGCGG--AAGCAGGCTTCTCTGCTCGGTGACAGTGTGCTGCTGAAACACAGG 641
Qy 1058 CTCAGCGCTTCCAGGGGATGTGNCAGCCCCCGGCTCAGCAGTGGGAAACACAGGCGCT 1117
Db 642 CTCAGCGCTTCCAGGG--ATCTGCCAGCCCCCGGCTCAGCAGTGGGAAACACAGGCGCT 698
Qy 1118 CGNAGCAGNAGNAGTGGGGGCAAGCAGAAATGCTCCAGAGATTTCANCTGAGC 1177
Db 699 CAGCAGCAAG-----GTGGGGGCAAGCAATGCTTCCAGGATTTTCACAC--TGA 748
Qy 1178 CCNTGCCCCANCCCTGTCTGAADAAAAAYTNCGCCACGTCGTAAGAGACAGAGGAGGATG 1237
Db 749 GCGCTGCCCCACCTGCTGAGAAACACT--CGCCAGCTGAAGAGACA----GAGGAGG 802
Qy 1238 GNCAGGATTTNACCTTGGGGGAAACAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTN 1297
Db 803 ATGGCAGGAGTTACCTCGGGGAAACAAACAGG-----ATCTTCTGCCCCCTGCTCAGT 856
Qy 1298 CGAGTTGCGCTGNACCCGCTTGGANTCAGTCACCATTTGTTGGCAGANCAGGGGAGAGCA 1357
Db 857 CGAGTTGCGCTGGA--CCGCTTGGATCAGTGGCCATTTGCTGGCAGAC--AGGGAGAGCA 912
Qy 1358 GCTTCCAGCCT--GGGTGAGAGGGGTGGGAGGCCCTTTCGCCCCCTCACCCCTNCCAGGCT 1416
Db 913 GCTTCCAGCCTGGGTCAAAAGGGGTGGGAGGCCCTTTCGCCCCCTCACCCCTCAGGCT 972
Qy 1417 GCTGTGAGAGTGTCAAGTGTGTAAAGGNCCTAAANTCTCAGNTTTCAGTGAGCAACAGG 1476
Db 973 GCTGTGAA--AGTNTCAAGTGTGTAAAGGCCCCAAACTC--AGGTTTCAAGTGAGCAACAGG 1028
Qy 1477 TNCAGCAGGTATGCCGCCGNTAGGTTAANNNGGGGGCCCTCTNAAACCCCTTGCCTTNGG 1536
Db 1029 TCAGCAGTATTGCGGGGCCCTTAGGTAA---GGGGGGGCCCTCTAANAACCCCTTGGCTGGC 1085
Qy 1537 CCTWCACCTNCGCAGCTCANCCCCCTTTTGGGTGTAGGGGAAAGAA 1583
Db 1086 CTAACGGGNCAGGTTACCC--CCCTTTTGGGGGAAAGGGGAAAGAA 1130
```

```
RESULT 9
LOCUS AL556130 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
DEFINITION AL556130 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DK011YM14 5-PRIME, mRNA sequence.
ACCESSION AL556130
VERSION AL556130.3 GI:45860849
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1133)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31277934.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8426.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DK011BG07QP1&c=8426.r.
FEATURES
Location/Qualifiers
1..1133
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YM14"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/primer="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized."
ORIGIN
Query Match 24.2%; Score 636.4; DB 1; Length 1133;
Best Local Similarity 84.7%; Pred. No. 1.le-154;
Matches 946; Conservative 6; Mismatches 115; Indels 50; Gaps 21;
Qy 824 AGCTTCAGCTTGTGTACCAAAAGACGGGACGCTGTTGCTTACGACATGGAGGTGCCAGA 883
Db 1 AGCTTCAGCTTGTGTACCAAAAGACGGGACGCTGTTGCTTACGACATGGAGGTGCCAGA 60
Qy 884 CTCGGGATCGACTGCGAGTGCACACTGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAG 943
Db 61 CTCGGGATCGACTGCGAGTGCACACTGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAG 120
Qy 944 GGTCAAGCATGGCAGCTGGAGAACAGGCCCTTACCTGCCCTCCACCGCGGTCCACA 1003
Db 121 GGTCAAGCATGGCAGCTGGAGAACAGGCCCTTACCTGCCCTCCACCGCGGTCCACA 180
Qy 1004 CTGCCGAAAGCAGCCTTCTGCTGCTGTCAGCATGTGCTGCCCTGAAAACACAGGCTCAGC 1063
Db 181 CTGCCG--AAGCAGCCTTCTGCTGCTGTCAGCATGTGCTGCCCTG--AAACACAGGCTCAGC 238
Qy 1064 GGTTCACAGGGATTTGNCACAGCCCCCGGCTCAGCAGTGGGAAACAGGCGCTCGNCAG 1123
Db 239 GGTTCACAGGG--ATCTGCCAGCCCCCGGCTCAGCAGTGGGAAACAGGCGCTCGCAGCAG 295
Qy 1124 CNAGCNAAGTNGGGGCAAGNAGATGCTCTCCAGGATTTTCACMCTCTGAGCCCTG 1183
Db 296 CAAG-----GTGGGGGCAAGCAGATGCTCTCCAGGATTTTCACCC---TGAGCCCTG 345
Qy 1184 CCCANCCCTGCTGAADAAAAACAYTNCGCCACGCTGAAGAGACAGAGGAGGATGGNCAGG 1243
```



Db	63	GGGACATGGGAGCCACATGCGCCCTCTGTCTCTCTCTGGATCTGCGCCCGCGCAGGAACCCCAACC	12
Qy	405	TCACCAGAGCGGAAAGCAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAATAAGATTA	464
Db	123	TCACCCGAGCGAAGGCAACCTTCTCTGAA CAGGAGCTACAGCAGCTGGAGATAGATTA	182
Qy	465	TTCTCTCAACAGCCTGTCTCCAGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTCTCTCGTGC	524
Db	183	TTTCTGAACAGCCTGTCTCCAGCCGTCTCTCCCTGGAGGAGCAGGAGCAGATCTCTCTCGTGC	242
Qy	525	CTCAGCATCGACAGCCTCTCCCTCTCGGATGACAGTGAGAGNACCCATCAAGAGCCTCT	584
Db	243	CTCAGCGTTCGACAGCCTCTCCCTCTCAGATGACAGTGAGAGAAACCCGTCCAAGGCTCTCT	302
Qy	585	CAAAAGCTCGGGGACACCTTGAGCTCAGGCGCTACACTCTCGAGCAGCAGGCGCGAGGCT	644
Db	303	CAGAGCTCGGGGACACACTGAGTTCGGCGTGCATTCTGTGGAGCAGCCAGGAGGCG	362
Qy	645	CGAAGCTTCAGCTGGAACATGGTCTGTGGCCGGGGGGCCACCGACACCCCAAGCTAT	704
Db	363	CGCAGCTCCAGCTGGAACATGGTCTGTGGCCGGGGGGCCACTGACACACCAAGCTAT	422
Qy	705	TTCAATGGTGTGAAGTTCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAG	764
Db	423	TTCAACGGTGTGAAGTTCAAATACAGTCTCTCAATGGCGAACACCTGCACATCCGGGAG	482
Qy	765	TTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCA	824
Db	483	TTCCACCGGGTCAAAGTGGGAGACATGCAACAGGCATCAGCAGCAGATCCAGCTGCA	542
Qy	825	GCCTTCAGCTTGGTTCACAAAGACGGGACGCTCTTTCGCTACGACATGAGGAGTGCAGAC	884
Db	543	GCCTTCAGCTTGGTGACAAAGACGGGACGCTCTTTCGCTACGACATGAGGAGTGCAGAC	602
Qy	885	TCGGGCATCACTGTCAGTCAGTCAGTCTGGCCCTGATGTCAGCTTCCCTCGAGCTGGAGG	944
Db	603	TCAGGCATCACTGTCAGTCAGTCTGGCCCTGATGTCAGCTTCCCTCGAGCTGGAGG	662
Qy	945	GTCAAGCATGGCGAGCTGGAGACAGGCGCCCTAAC	979
Db	663	GTCAAGCATGGCGAGCTGGAGACAGGCGCCCTAAC	697

RESULT 12

Bi161002

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Bi161002

602845208F1 NIH\_MGC\_42 Homo sapiens CDNA clone IMAGE:5019481 5', mRNA sequence.

Bi161002

Bi161002.1 GI:14621003

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 806)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: L16M1834 row: 0 column: 02

High quality sequence stop: 600.

Location/Qualifiers

1..806

organism="Homo sapiens"

FEATURES

source

```
RESULT 13
BP267282
LOCUS
DEFINITION
580 bp mRNA linear EST 16-SEP-2004
Sugano cDNA library, Thyroid JTH Homo sapiens cDNA clone
JTH06618, mRNA sequence.
ACCESSION
BP267282.1 GI:52182514
VERSION
BP267282.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 580)
Suzuki Y., Yamashita R., Shirota M., Sakakibara Y., Chiba J.,
Mizushima-Sugano J., Nakai K. and Sugano S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL
Department of Virology
COMMENT
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukui@ims.u-tokyo.ac.jp.
FEATURES
Location/Qualifiers
source
1..580
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH06618"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/note="thyroid tumor"
ORIGIN
Query Match 21.4%; Score 564.2; DB 5; Length 580;
Best Local Similarity 99.3%; Pred. No. 6.8e-136; Indels 1; Gaps 1;
Matches 577; Conservative 0; Mismatches 3;
QY 461 ATTATTTCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGAAATTTCTTC 520
Db 1 ATTATTTCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGAAATTTCTTC 60
QY 521 GTGCTCAGCATCGACAGCCTCTCCCTGTCGATGACAGTGAAGACCATCAAGGC 580
Db 61 GTGCTCAGCATCGACAGCCTCTCCCTGTCGATGACAGTGAAGACCATCAAGGC 120
QY 581 CTCTCAAGCTCGCGGACACCTCAGCTCAGGCGTACACTCTGGAGCAGCAGGCCGA 640
Db 121 CTCTCAAGCTCGCGGACACCTCAGCTCAGGCGTACACTCTGGAGCAGCAGGCCGA 180
QY 641 GGCTCGAAGCTCCAGCTGGAAACATGGTGTGGCCCGGGGGCCCGCCAGCCCAAG 700
Db 181 GGCTCGAAGCTCCAGCTGGAAACATGGTGTGGCCCGGGGGCCCGCCAGCCCAAG 240
QY 701 CTATTTCAATGTGTGAAGTCCAAATACAGTCTCTTAATGTGAACACCTGCACATCG 760
Db 241 CTATTTCAATGTGTGAAGTCCAAATACAGTCTCTTAATGTGAACACCTGCACATCG 300
QY 761 GGAGTTCCACCGGGTCAAAAGTGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGC 820
Db 301 GGAGTTCCACCGGGTCAAAAGTGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGC 360
QY 821 TGCAGCTTCAGCTTGGTCAACCAAGACCGGAGCTGTTCCTAGACATGAGGTGC 880
Db 361 TGCAGCTTCAGCTTGGTCAACCAAGACCGGAGCTGTTCCTAGACATGAGGTGC 420
QY 881 AGACTCGGCATCGACCTGCAGTGCACACTGGGCCCTGTATGGCAGCTTCCTTGGAGTGG 940
Db 421 AGACTCGGCATCGACCTGCAGTGCACACTGGGCCCTGTATGGCAGCTTCCTTGGAGTGG 480
QY 941 GAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCC 1000
```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5019481"
/tissue_type="epithelioid carcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_42"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"
ORIGIN
Query Match 21.6%; Score 568.2; DB 4; Length 806;
Best Local Similarity 91.9%; Pred. No. 6.8e-137; Indels 14; Gaps 9;
Matches 701; Conservative 0; Mismatches 48;
QY 1 CCCTCTTCACAGCCAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGGTGT 60
Db 42 CCCTCTTCACAGCCAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGGTGT 101
QY 61 CTCAGCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGA 120
Db 102 CTCAGCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGCTCTGAAGA 161
QY 121 GCCTTTGGAGGGAGATATAAGAACCAAGACATCCAGCCCAATCAAGCAATATACC 180
Db 162 GCCTTTGGAGGGAGATATAAGAACCAAGACATCCAGCCCAATCAAGCAATATACC 221
QY 181 ACCAGACCTCTCCATGCCAGCCGAGAGCTTTGCGCAAGGGCCCCAGGGCCCGCCAG 240
Db 222 ACCAGACCTCTCCATGCCAGCCGAGAGCTTTGCGCAAGGGCCCCAGGGCCCGCCAG 281
QY 241 CTGAGGAGACACAGGAGAGCCCTTAAGCTCAGCTCTCTC-TCCACACAGAGCCCA 299
Db 282 CTGAGGAGACACAGGAGAGCCCTTAAGCTCAGCTCTCTC-TCCACAGAGCCCA 341
QY 300 GAGCAACCAAGTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCC 359
Db 342 GAGCAACCAAGTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAAGCC 401
QY 360 TTACTCTGTCTCTCTGAGGCA-GCCCTTGCAGAAACCCAGCTCACCAGCGGA 418
Db 402 TTACTCTGTCTCTCTGAGGCAATGCCCTTCGAGAAAGCCAGCTCACCAGCGGA 461
QY 419 AGCAACCTCTCCGAGCAGGAATGCGAGAGCTGGAAATAGAAATTTCTCTCAACAGCCT 478
Db 462 AGCAACCTCTCCGAGCAGGAATGCGAGAGCTGGAAATAGAAATTTCTCTCAACAGCCT 521
QY 479 GTCCAGCATTCTCTCTGG-AGAGAGGAGCAAAATTTCTCTGTGCTCAGCATCGACA 537
Db 522 GTCCAGCATTCTCTCTGGAGGAGAGGCAAAATTTCTCTGTGCTCAGCATCGACA 581
QY 538 GC-CTCTCTCTGTGGATGACAGTGAAGAA-CCCATCAAGGCTCTCAAGCTCGCG 595
Db 582 GCATCTCTCTGTGGATGACCGTGAAGAAACCCCATTTAGAGCTCTCAAGTCCGCA 641
QY 596 GG---ACACCTCTGAGCTCAGGCG---TACACTCTCTGGAGCAGCCAGGCGAGCTCGAA 648
Db 642 GCGCAACCTTGTAGCTCAGGCGGTACAACCTCTCTGGAGCAACCAAGGCGAGGTTCGAA 701
QY 649 GCTCA-GETGAAACATGGTGTGGCCCGGGGGCCCGCCAGCAGCCCAAGCTATTTC 707
Db 702 GCTCCAGCTGACCTGTGTGTGGCCCGGGGGAGGCAACCGAAACCCCAAGCTATTTC 761
QY 708 AATGTGTGAAGTCCAAATACAG-TCTTTAATGGTGAACAC 749
Db 762 AATGTGTGAAGTCCAAATACAGCCCTTACTTGTGAACAC 804
```

```
Db 481 GAGGTCAAGCATGCCAGCTTGGAAACAGGCCCTTAAACCTGCTCCACCGCGGCTCC 540
Qy 1001 ACATGCGGGAAGCAGCTTCTCTCTCGGTGACGATGCT 1041
Db 541 ACATGCGGG-AAGCAGCTTCTCTCTCGGTGACGATGCT 580

RESULT 14
AL578455/c
LOCUS AL578455 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens linear EST 07-APR-2004
DEFINITION CDNA clone CS0DK011Y14 3-PRIME, mRNA sequence.
ACCESSION AL578455
VERSION AL578455.3 GI:46257344
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1095)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:31316669.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8426.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/s=CS0DK011BG07NP1&c=8426.r.
FEATURES
source
1. 1095
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011Y14"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 21.0%; Score 553; DB 1; Length 1095;
Best Local Similarity 82.2%; Pred. No. 7e-133;
Matches 899; Conservative 9; Mismatches 129; Indels 57; Gaps 25;
Qy 1458 GNTTCAGTGCAGAACCCAGTTCAGGATGCGCCGCGGCTAGGTTAANNNGGCGGCGCT 1517
Db 1038 GGTTCAGTGCAGAACCCAGTTCAGGATGCGCCGCGGCTAGGTTAANNNGGCGGCGCT 987
Qy 1518 CTNAAACCCCTTGCTNGGCTNCCACCTNGGCCAGCTCANCCCTTTTGGGTGTAGGGGA 1577
Db 986 YCTTAACCCCTTGCTNGGCTNCCACCTNGGCCAGCTCANCCCTTTTGGGTGTAGGGGA 931
Qy 1578 AAGNATGCTGACCTGGGAAGGCTGCCCTGGTGAATACACACACTTTTCAGGTGT 1637
Db 930 AAGAATG-CTGACCTGGGAAGGCT-CCCTGGTGAATACACACACTTTTCAGGTGT 873
Qy 1638 TGCAACACAGTCTCTGAGTTCAGCTCTGTTTCAGCAGGACCAAGAGTGTGAAGT 1697
Db 872 TGCAACACAGTCTCTGAGTTCAGCTCTGTTTCAGCAGGACCAAGAGTGTGAAGT 813
Qy 1698 GAAGTGGTTCTCAGTNCCCAAGACATGTGCCCTTTTGTGCTGCTACCACTCTTCCCA 1757
```

```
Db 812 GAAGTGGTTCTCAGT-CCCCAGACATGTGCCCTTTTGTGCTGCTACCACTTCCCCA 754
Qy 1758 GAGCAGAGGCGCCGAGGCCCTTTCAGGCCAGCACTGCCCCAGAG-CTGCTGGCACTCAG 1816
Db 753 GAGCAGAGGCGCCGAGGCCCTTTCAGGCCAGCACTGCCCCAGAGCTNGTGGNACTCAG 694
Qy 1817 TTCCCTCATCTGTAAAGGTGAAGGTGATGATGATGATGATGATGATGATGATGATGATG 1876
Db 693 TTCCCTCATCTGTAAAGGTGAAGGTGATGATGATGATGATGATGATGATGATGATGATG 634
Qy 1877 ATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1936
Db 633 ATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 576
Qy 1937 CCACTNATCAGTGTNCCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1996
Db 575 CACTATCAGTGT---CCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 527
Qy 1997 CTGACATNCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2056
Db 526 NACTGANACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470
Qy 2057 ACNTCTGATGNACCTCAAGCACTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2116
Db 469 AC-TCTGATG-ACCTCAAGGACTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 412
Qy 2117 GGCAGTGCACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2176
Db 411 TGACACT-----GTAGGACATGATGATGATGATGATGATGATGATGATGATGATGATG 364
Qy 2177 GTCTTGAGCTGTGCANCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2236
Db 363 GTCTTGAGCTGTGCA-CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 309
Qy 2237 GAATAGCANCCCTTNCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2296
Db 308 GAATAGCACCCCTTNCCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
Qy 2297 CAGGACTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2356
Db 251 CAGAACTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 193
Qy 2357 CCTCTGGGCGAGGCGCGGAGACTCCGTTGGGAGAGTCTCATTTCCACATCTTGGCAA 2416
Db 192 CCTCTGGGCGAGGCGCGGAGACTCCGTTGGGAGAGTCTCATTTCCACATCTTGGCAA 133
Qy 2417 GACAGCTTTTNGTCCAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATG 2476
Db 132 GACAGCTTTT-GTCCAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATG 74
Qy 2477 CCCCAGCACATAAAGAACTGCGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2536
Db 73 CCCCAGCACATAAAGAACTGCGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 15
Qy 2537 TTGTAAGCAATAA 2550
Db 14 TAAAGNNAARAAA 1

RESULT 15
BU541005
LOCUS AGENCOURT_10326204 NIH_MGC_42 Homo sapiens cDNA clone IMAGE:6572558
DEFINITION 5', mRNA sequence.
ACCESSION BU541005
VERSION BU541005.1 GI:22851446
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 955)
```



AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLM2764 row: n column: 14  
 High quality sequence start: 31  
 High quality sequence stop: 630.  
 Location/Qualifiers  
 1..955  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6572558"  
 /tissue\_type="epithelioid carcinoma cell line"  
 /lab\_host="pH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_42"  
 /note="Organ: pancreas; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library. |"  
 ORIGIN

Query Match 19.58; Score 514.2; DB 5; Length 955;  
 Best Local Similarity 82.68; Pred No. 9.3e-123;  
 Matches 795; Conservative 1; Mismatches 122; Indels 45; Gaps 18;  
 QY 1622 ACACCTTTTCAGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGGTTACGCCAAGACCA 1681  
 DB 32 ACACCTTTTCAGTTGTTGCAACACAGGTCCTGAGTTGACCTCTGGTTACGCCAAGACCA 91  
 QY 1682 AAGAAGTGTGTAGTCAAGTGTCTCAGTNCACAGACATGTCGCCCTTTCTGCTG 1741  
 DB 92 AAGAAGTGTGTAGTCAAGTGTCTCAGT -CCCCAGACATGTGCCCTTTCTGCTG 150  
 QY 1742 CTACCACTTCTCCACAGCAGCAGGCCCCGAGCCCTTCAGGCCAGCAGCTGCCCCAGA 1801  
 DB 151 CTACCACTTCTCCACAGCAGCAGGCCCCGAGCCCTTCAGGCCAGCAGCTGCCCCAGA 210  
 QY 1802 CTCGCTGGCACTCAGTTCCCTCTCTGTAAAGTGAAGGTGATGCGAGGATATGCTGAC 1861  
 DB 211 CTCGCTGGCACTCAGTTCCCTCTCTGTAAAGTGAAGGTGATGCGAGGATATGCTGAC 270  
 QY 1862 AGGACAGTCTGTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921  
 DB 271 AGGACAGTCTGTGATGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 328  
 QY 1922 CCGGCGCCCCAGNCCCCTNATCATGTGTCNACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1981  
 DB 329 CCGGCGCCCCAGNCCCCTNATCATGTGTCNACGCTGCTGCTGCTGCTGCTGCTGCTGCT 377  
 QY 1982 CTNACAGNATCAGACATGACATNACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041  
 DB 378 CACAGCTCAGCATCAGCTGACATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 435  
 QY 2042 CGNACGGCACTTTGACATCTGATGACCTCAAGCACTTTTCAAGGCTGCTGCTGCTGCTGCT 2101  
 DB 436 CG-ACGGCACTTTGAC -TCTGATG-ACCTCAAGCACTTTTCAAGGCTGCTGCTGCTGCTG 492  
 QY 2102 CAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAG 2161

DB 493 GCAGGGTAGGCGACGTGACACT-----GTAGGAGCATAGCAACCCAGGAGATGGG 542  
 QY 2162 GTGNAAGGANCACAGTCTTGGAGCTGCCANCATCATGTGACTTNCCTCAAAACCTCTTNN 2221  
 DB 543 GTCAA--GGGACACAGTCTTGGAGCTGTCCA-CATGATGTGACT--CTCAAAACCTCTT-- 596  
 QY 2222 CCAGNATTTCTCTAAGAATAGCANGCCCTTNCCTTCCCATTTGCCAGCTTAGCTCTTCTC 2281  
 DB 597 -CCAGATTTCTTAAGAATAGCACCCCC--TTCCCATTTGCCAGCTTAGCTCTTCTC 653  
 QY 2282 CCAGGGAGGCTANTCTCAGGACTCAGTAGCATTAATCAGCTGTGNAATCGTCAGGGGT 2341  
 DB 654 CCAGGGAGGCTA-CTCAGGACTCAGTAGCATTAATCAGCTGTGNAATCGTCAGGGGT 712  
 QY 2342 GTCTGTAGCTCAACCTCTCTGGGCGAGGGGACGCCGAGACTCCGTTGGGAGAGCTCAAT 2401  
 DB 713 GTCTGTAGCTCAACCTCTCTGGGCGAGGGGACGCCGAGACTCCGTTGGGAGAGCTCAAT 772  
 QY 2402 CCACATCTTTGCCAAGACAGCCTTTTNGTCCAGCTGTCCATGTGATGATGATGATGATGAT 2461  
 DB 773 CCACATCTTTGCCAAGACAGCCTTTT-GTCCAGCTGTCCACATGATGATGATGATGATGAT 831  
 QY 2462 GGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGAGCCTTTGG----TACTGCAGAGT 2517  
 DB 832 GGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGAGCCTTTGGGTTGGTACTGCAAAAGTCT 891  
 QY 2518 CTGGGTTGTAGAGAACTC-TTTGTAAGCAATAAAGTTTGGGTTGGGTTGGTACTGCAAAAGTCT 2576  
 DB 892 GGGGTTGTAGAGAACTCTTTTGTAGCCACTAAAGTTGGGGTGTATCAACAATGTTAAA 951  
 QY 2577 AAA 2579  
 DB 952 AAA 954

Search completed: June 10, 2005, 20:27:33  
 Job time : 5555.69 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:18:24 ; Search time 5419.24 Seconds  
(without alignments)  
17042.190 Million cell updates/sec

Title: US-09-155-676B-1  
Perfect score: 1906  
Sequence: 1 cattgggtacgcgggtggcg.....tcattcttgataaaaaaaa 1906

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1899	99.6	1906	6	A66645 Sequence 1
2	1742	91.4	1782	6	AX089604 Sequence
3	1555	81.6	2873	6	AX089606 Sequence
4	1555	81.6	3139	6	AX089605 Sequence
5	1138.2	59.7	2248	6	BD157657 Primer fo
6	1138.2	59.7	2248	6	AX879111 Sequence
7	1138.2	59.7	2248	9	AK022425 Sequence
8	909.4	47.7	1033	6	CQ714292 Sequence
9	794.4	41.7	5833	9	HS0808428
10	792.8	41.6	6045	6	BD160227 Primer fo
11	792.8	41.6	6045	6	AX883269 Sequence
12	792.8	41.6	6045	9	AK023827 Homo sapi
13	672.2	35.3	730	6	BD146965 Primer fo
14	672.2	35.3	730	6	AX866903 Sequence
15	630.8	33.1	1229	9	BC041583 Homo sapi
16	376	19.7	161864	9	AC007601 Homo sapi
17	376	19.7	174136	2	AC126766 Homo sapi
18	376	19.7	198521	2	AC145723 Homo sapi
19	376	19.7	201237	2	AC141081 Homo sapi

20	376	19.7	202689	2	AC034281 Homo sapi
21	361.8	19.0	198703	2	AC144590 Gorilla g
22	353.8	18.6	203216	2	AC097328 Pan trogl
23	352.2	18.5	132393	2	AC138881 Homo sapi
24	352.2	18.5	140011	2	AC008864 Homo sapi
25	352.2	18.5	158388	9	AC008740 Homo sapi
26	352.2	18.5	168288	9	AC009124 Homo sapi
27	352.2	18.5	171940	9	AC025279 Homo sapi
28	352.2	18.5	174010	2	AC023814 Homo sapi
29	352.2	18.5	178081	2	AC023463 Homo sapi
30	352.2	18.5	194943	2	AC142201 Homo sapi
31	352.2	18.5	200812	9	AC133555 Homo sapi
32	352.2	18.5	233479	2	AC126758 Homo sapi
33	350.6	18.4	156989	2	AC069176 Homo sapi
34	293.4	15.4	1431	10	AF399755 Mus muscu
35	263	13.8	5496	9	AK074072 Homo sapi
36	242.4	12.7	198521	2	AC145723 Homo sapi
37	242.4	12.7	202689	2	AC034281 Homo sapi
38	242.2	12.7	203216	2	AC097328 Pan trogl
39	233.6	12.3	148230	2	AC141548 Rattus no
40	233.6	12.3	165881	2	AC133732 Rattus no
41	233.6	12.3	263443	2	AC121644 Rattus no
42	221	11.6	177955	2	AC131034 Mus muscu
43	215.8	11.3	156989	2	AC069176 Homo sapi
44	205.4	10.8	227	6	CQ731480 Sequence
45	193	10.1	367	6	CQ750241 Sequence

ALIGNMENTS

RESULT 1	A66645	Sequence 1 from Patent WO9737016.	1906 bp	DNA	linear	PAT 29-MAR-1999
LOCUS	A66645					
DEFINITION	A66645					
ACCESSION	A66645					
VERSION	A66645.1	GI:4538138				
KEYWORDS	unidentified					
SOURCE	unclassified					
ORGANISM	unclassified					
REFERENCE	1 (bases 1 to 1906)					
AUTHORS	Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.					
TITLE	MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR PREPARATION AND USE					
JOURNAL	Patent: WO 9737016-A 1 09-OCT-1997;					
COMMENT	VEDA RES & DEV (IL)					
FEATURES	Other publication AU 2175597 19971022.					
source	Location/Qualifiers					
	1. .1906					
	/organism="unidentified"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32644"					

ORIGIN

Query Match	99.6%	Score 1899;	DB 6;	Length 1906;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1906;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	CATTGGGTACCGCGTGGCGGCTCTAGAAATAGTGAATCCCGGGCTCGAGGAATCG	60	
Db	1	CATTGGGTACCGCGTGGCGGCTCTAGAAATAGTGAATCCCGGGCTCGAGGAATCG	60	
Qy	61	ATTTCGAGGCCACGAAGCGCGCGCGCANGCACCGCGCGCGGANGAGGNCATG	120	
Db	61	ATTTCGAGGCCACGAAGCGCGCGCGCANGCACCGCGCGCGGANGAGGNCATG	120	
Qy	121	ACCGGATCNCAGAACNATGACAAAAGCAATTTCTGCTGAGCGACTGCTGGATGCAATG	180	
Db	121	ACCGGATCNCAGAACNATGACAAAAGCAATTTCTGCTGAGCGACTGCTGGATGCAATG	180	
Qy	181	AAACAGTGCAGATCCGCTTTGAGGAGAAAGAGATTCCCTCGGATTCGACAGCAGG	240	
Db	181	AAACAGTGCAGATCCGCTTTGAGGAGAAAGAGATTCCCTCGGATTCGACAGCAGG	240	

181 AAA CAGTCCAGATCCGCTTTTNGAGGAGAAAGAGATTGCTCGGATTCGACAGCAGG 240  
241 GTCACTGTCTGTGTGTCGCCAGTTTGAAGCCCTCTCTGACGATATGCTTGAAGAGAGTCTGA 300  
241 GTCACTGTCTGTGTGTCGCCAGTTTGAAGCCCTCTCTGACGATATGCTTGAAGAGAGTCTGA 300  
301 GGATTTGGACACTCAGAGCGGACGATCAAGCAGAGCAGCGGGCTTTGGCAGCAAAACCGAA 360  
301 GGATTTGGACACTCAGAGCGGACGATCAAGCAGAGCAGCGGGCTTTGGCAGCAAAACCGAA 360  
361 ACAGAGCCCGTGTCTTGTGTAAGAGAGTCTCTCAACAGACGAGCTCGACGC 420  
361 ACAGAGCCCGTGTCTTGTGTAAGAGAGTCTCTCAACAGACGAGCTCGACGC 420  
421 TTCTACTCCCTGGGCGACATCGCTCAGACGTGGGCGGGGTGCGGCTTGGCTGCGTGT 480  
421 TTCTACTCCCTGGGCGACATCGCTCAGACGTGGGCGGGGTGCGGCTTGGCTGCGTGT 480  
481 GCCCTCAACGAACACTCCCTGGAGCGCTTACTCTGACATGCTCTCTGGGCGACCGCTCGAGG 540  
481 GCCCTCAACGAACACTCCCTGGAGCGCTTACTCTGACATGCTCTCTGGGCGACCGCTCGAGG 540  
541 CTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTT 600  
541 CTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTT 600  
601 CCTACCATGGCAGCAGGTCTGAATCCCATACTCTTTGGGATTAACATCGACAAACAGGAT 660  
601 CCTACCATGGCAGCAGGTCTGAATCCCATACTCTTTGGGATTAACATCGACAAACAGGAT 660  
661 TTGAACGGGCGAGAGTAAAGTTTGTCTCCACCGCTTTCAGACCTCTTAAAGGAGTCAACGCAG 720  
661 TTGAACGGGCGAGAGTAAAGTTTGTCTCCACCGCTTTCAGACCTCTTAAAGGAGTCAACGCAG 720  
721 AACGTGACCTCTTGTCTCAAGAGTCCACGCAAGAGTGTGAGCAGCCTTTTCAGGGAGATC 780  
721 AACGTGACCTCTTGTCTCAAGAGTCCACGCAAGAGTGTGAGCAGCCTTTTCAGGGAGATC 780  
781 ACAGCTCTCTGCGCTCTCCATCTCATCAAACTGAACTGAACTGAACTGAACTGAACTGAACTG 840  
781 ACAGCTCTCTGCGCTCTCCATCTCATCAAACTGAACTGAACTGAACTGAACTGAACTGAACTG 840  
841 TCGTCTCAGGAATCTCAGTGTCTGATGCAATGCAAAAGAGAGCGGAGAGGAGGAGGAGGAGGAG 900  
841 TCGTCTCAGGAATCTCAGTGTCTGATGCAATGCAAAAGAGAGCGGAGAGGAGGAGGAGGAGGAG 900  
901 AAGTGACCAACATAATCTCAATTTGATGATGAGGAGAGATGAGCAGAACTCTGGGAGACGTGT 960  
901 AAGTGACCAACATAATCTCAATTTGATGATGAGGAGAGATGAGCAGAACTCTGGGAGACGTGT 960  
961 TTAAAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA 1020  
961 TTAAAAAGACACCTGGGCGAGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA 1020  
1021 ATATCATGTCCGCTTTGAAAGCCCTTTGGGCTTAATCCAAATGGAATCAGAGCAGCAA 1080  
1021 ATATCATGTCCGCTTTGAAAGCCCTTTGGGCTTAATCCAAATGGAATCAGAGCAGCAA 1080  
1081 CTATCGGAAATGATTCCTGTCTTTGAA CGGGAGTTTGGGTACCAAGAGCTTGATGT 1140  
1081 CTATCGGAAATGATTCCTGTCTTTTGAACGGGAGTTTGGGTACCAAGAGCTTGATGT 1140  
1141 GAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATACGCTGATGAAACTCATC 1200  
1141 GAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATACGCTGATGAAACTCATC 1200  
1201 AGGAAGGAGACACAGGGCCACTCGAGTTCGCGGAGAGCACTTGAAGGGGAAACACCTG 1260  
1201 AGGAAGGAGACACAGGGCCACTCGAGTTCGCGGAGAGCACTTGAAGGGGAAACACCTG 1260  
1261 CCTCTCCAGATGACACGTGGGCTCGCTGAAGTGTGTCACATGACTCCGACATCCT 1320  
1261 CCTCTCCAGATGACACGTGGGCTCGCTGAAGTGTGTCACATGACTCCGACATCCT 1320

1321 CTTCCCTGTCTCAGTGGCTGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAAGCTTGGGA 1380  
1321 CTTCCCTGTCTCAGTGGCTGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAAGCTTGGGA 1380  
1381 GAAACGGACAGGACACAGAGACCACTTTCTCCGGATCTCTGGGACTTCGGTACAGTGTGGA 1440  
1381 GAAACGGACAGGACACAGAGACCACTTTCTCCGGATCTCTGGGACTTCGGTACAGTGTGGA 1440  
1441 AGCCAGCTCTCAGGGCCACGGAAGTCTCTGAGCAGCCTGTACTTCTGCTCAGTGCCA 1500  
1441 AGCCAGCTCTCAGGGCCACGGAAGTCTCTGAGCAGCCTGTACTTCTGCTCAGTGCCA 1500  
1501 GAGTCCATGACAAATTAGTGAATCGCCAGGCACTGTGGCCATGATGAACAGGAAGAT 1560  
1501 GAGTCCATGACAAATTAGTGAATCGCCAGGCACTGTGGCCATGATGAACAGGAAGAT 1560  
1561 GAGTGGAGGAGGAGAAACAGATCACTGCGAAAACCTGCTCGACGCTGAGATGAGCACTCA 1620  
1561 GAGTGGAGGAGGAGAAACAGATCACTGCGAAAACCTGCTCGACGCTGAGATGAGCACTCA 1620  
1621 GCCGCGCTCCGCAAGAGGTGGACACTTGAAGAGGAGGTGGCTGAACAGGAGGAGCGG 1680  
1621 GCCGCGCTCCGCAAGAGGTGGACACTTGAAGAGGAGGTGGCTGAACAGGAGGAGCGG 1680  
1681 CAGGCGATGAAGTCCAGGCGCTGGCCAGCTATCTTTGTCTATTTTGTGAGGAGATCTTAA 1740  
1681 CAGGCGATGAAGTCCAGGCGCTGGCCAGCTATCTTTGTCTATTTTGTGAGGAGATCTTAA 1740  
1741 CCCACGTGAGAACCATGTGTGGAGAAATGAGGAGGAGAGAGAAATCCAAACAGTTCCTGA 1800  
1741 CCCACGTGAGAACCATGTGTGGAGAAATGAGGAGGAGAGAGAAATCCAAACAGTTCCTGA 1800  
1801 TAGTCTCATTTGAGTCTCTGATCCAGTCTTCTGAGGCTGTGTTCTCTGAGCTTTT 1860  
1801 TAGTCTCATTTGAGTCTCTGATCCAGTCTTCTGAGGCTGTGTTCTCTGAGCTTTT 1860  
1861 CATGTATGTGAGCCATAAATTTGCTTTTCATTCCTTGAAAAAAA 1906  
1861 CATGTATGTGAGCCATAAATTTGCTTTTCATTCCTTGAAAAAAA 1906

RESULT 2  
AX089604  
LOCUS Sequence 4 from Patent WO0116314.  
DEFINITION AX089604  
ACCESSION AX089604  
VERSION AX089604.1 GI:13443796  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.  
TITLE Iren protein, its preparation and use  
JOURNAL Patent: WO 0116314-A 4 08-MAR-2001;  
YEDA RESEARCH AND DEVELOPMENT Co. LTD. (IL)  
FEATURES  
source Location/Qualifiers  
1..1782  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 91.4%; Score 1742; DB 6; Length 1782;  
Best local Similarity 99.6%; Pred. No. 0;  
Matches 175; Conservative 0; Mismatches 4; Indels 3; Gaps 3;  
Qy 118 ATGAGCGGATCNCNGAACNATGACAAAAGACAAATTTCTGTGGAGCGACTCTCGATGCA 177  
Db 1 ATGAGCGGATCAGAGAACAAATGACAAAAGACAAATTTCTGCTGGAGCGACTCTCGATGCA 60

178 GTGAAACAGTGCAGATCCGCTTTNGAGGAGAAAGAGATTGCTCGGATTCGACAGC 237  
Db |||||  
61 GTGAAACAGTGCAGATCCGCTTTGAGGAGAAAGAGATTGCTCGGATTCGACAGC 120  
Qy |||||  
238 AGGTCACTGTCTGTGTCGCCAGTTTGAAGCCGTCTCTGAGCATGCTTTGAAGAGAGT 297  
Db |||||  
121 AGGTCACTGTCTGTGTCGCCAGTTTGAAGCCGTCTCTGAGCATGCTTTGAAGAGAGT 180  
Qy |||||  
298 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGACGCGGCTTTGCGAGCAAAACC 357  
Db |||||  
181 CGAGGATTGGCACTCACAGCGGCAGCGATCAAGCAGGACGCGGCTTTGCGAGCAAAACC 240  
Qy |||||  
358 GAAACAGAGCCGTGTTCTGTGTAAGTCTGAGAGAGTCTCAACAGAGCAGAGCTGCAG 417  
Db |||||  
241 GAAACAGAGCCGTGTTCTGTGTAAGTCTGAGAGAGTCTCAACAGAGCAGAGCTGCAG 300  
Qy |||||  
418 CGCTTCTACTCCCTGCGCCACATCGCTCAGAGTGGCGCGGGTCCGCGCTGGCTGCGC 477  
Db |||||  
301 CGCTTCTACTCCCTGCGCCACATCGCTCAGAGTGGCGCGGGTCCGCGCTGGCTGCGC 360  
Qy |||||  
478 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGCTCTGGCGGACCGCTGC 537  
Db |||||  
361 TGTGCCCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGCTCTGGCGGACCGCTGC 420  
Qy |||||  
538 AGGTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGGTCAAGTATG 597  
Db |||||  
421 AGGTGAGCACTTTTATGAAGACTGTCTTTTGTGATGATGAAGAGGTCAAGTATG 480  
Qy |||||  
598 CTTCCTACCATGGCAGAGTCTGAACCTCCATCTCTTCGGATTAACTCGACACAAG 657  
Db |||||  
481 CTTCCTACCATGGCAGAGTCTGAACCTCCATCTCTTCGGATTAACTCGACACAAG 540  
Qy |||||  
658 GATTGAAACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACG 717  
Db |||||  
541 GATTGAAACGGGAGAGTAAGTTTGTCTCCACCGTTTTCAGACCTCTTAAAGAGTCAACG 600  
Qy |||||  
718 CAGAACGTGACCTCTCTGCTGAAAGAGTCAACGCAAGAGTGAAGAGCTGTTCAGGGAG 777  
Db |||||  
601 CAGAACGTGACCTCTCTGCTGAAAGAGTCAACGCAAGAGTGAAGAGCTGTTCAGGGAG 660  
Qy |||||  
778 ATACAGCCTCTCTGCGCTCTCCATCTCATCAACCTGAAACAGGACCGA-CCCTTG 836  
Db |||||  
661 ATACAGCCTCTCTGCGCTCTCCATCTCATCAACCTGAAACAGGACCGA-CCCTTG 720  
Qy |||||  
837 CCTGTCTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGAGCGGAAGAGAAA 896  
Db |||||  
721 CCTGTCTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGAGCGGAAGAGAAA 780  
Qy |||||  
897 AAGAAAGTGACCAACATAATCTCATTTTGTATGATGAGGAAGATGAGCAGAACTCTGGGGAC 956  
Db |||||  
781 AAGAAAGTGACCAACATAATCTCATTTTGTATGATGAGGAAGATGAGCAGAACTCTGGGGAC 840  
Qy |||||  
957 GTGTTTAAAGACACCTGGGCGAGGAGCTCAGAGACAACTCCGACCGCTCTCT 1016  
Db |||||  
841 GTGTTTAAAGACACCTGGGCGAGGAGCTCAGAGACAACTCCGACCGCTCTCT 900  
Qy |||||  
1017 GTCAATATCATGTCCGCTTTGAAAGCCCTTGGGCTTAATCTCAATGGAA-TCAGAGC 1075  
Db |||||  
901 GTCAATATCATGTCCGCTTTGAAAGCCCTTGGGCTTAATCTCAATGGAAAGTCAGAGC 960  
Qy |||||  
1076 AGCAACTCATGGAAATTTGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT 1135  
Db |||||  
961 AGCAACTCATGGAAATTTGATTCCTGTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTT 1020  
Qy |||||  
1136 GATGTGAAAGCATCGATGATGAAGATGTGATGAAACCAAGATGACGTGTATGGAAC 1195  
Db |||||  
1021 GATGTGAAAGCATCGATGATGAAGATGTGATGAAACCAAGATGACGTGTATGGAAC 1080  
Qy |||||  
1196 TCATCAGGAAGGACACAGGCGCACTCGAGTTCGCCCGAGAGCCACTGGAGGGAAC 1255  
Db |||||  
1081 TCATCAGGAAGGACACAGGCGCACTCGAGTTCGCCCGAGAGCCACTGGAGGGAAC 1140  
Qy |||||  
1256 ACCTGCCTCTCCAGATGCACAGCTGGGCTCCGCTGAAGGTGTGCACAATGACTCCGAC 1315

1141 ACCTGCCTCTCCAGATGCACAGCTGGGTCCGCTGAAGGTGTCACAAATGACTCCGAC 1200  
Qy |||||  
1316 ATCTCTTCTCCCTGTAGTGGCGTCTCTACAGCCAGCAGATGCCCTCGGAAGC 1375  
Db |||||  
1201 ATCTCTTCTCCCTGTAGTGGCGTCTCTACAGCCAGCAGATGCCCTCGGAAGC 1260  
Qy |||||  
1376 CTGGAAACGGGACAGGACAGGACCAAGTCTCTCCGATCCTGGACTTCGGTACAGT 1435  
Db |||||  
1261 CTGGAAACGGGACAGGACAGGACCAAGAGCACAGTCTCTCCGATCCTGGACTTCGGTACAGT 1320  
Qy |||||  
1436 GTGGAAGCAGCTCTCCAGCCAGGAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1494  
Db |||||  
1321 GTGGAAGCAGCTCTCCAGCCAGGAGTCTCTGAGCAGCCTGTTACCTTCTGCTCA 1380  
Qy |||||  
1495 GTGCCAGAGTCCATGAACTTAGTGAATGCGCAGGACCACTGTGGCCATGATGAACAGG 1554  
Db |||||  
1381 GTGCCAGAGTCCATGAACTTAGTGAATGCGCAGGACCACTGTGGCCATGATGAACAGG 1440  
Qy |||||  
1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGAAACCTGTGACGCTGAGATGAG 1614  
Db |||||  
1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGAAACCTGTGACGCTGAGATGAG 1500  
Qy |||||  
1615 CACTCAGCCGCTCCGCGCAGAGTGGACACCTTGAAAGGAGGCTGGCTGAACAGAG 1674  
Db |||||  
1501 CACTCAGCCGCTCCGCGCAGAGTGGACACCTTGAAAGGAGGCTGGCTGAACAGAG 1560  
Qy |||||  
1675 GAGCGCAGGCGCATGAAGTCCAGGCGTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1734  
Db |||||  
1561 GAGCGCAGGCGCATGAAGTCCAGGCGTGGCCAGCTATCTTTGCTATTTTGTGAGGAGA 1620  
Qy |||||  
1735 TTCTAACCCACGTGAGAACCATGTGTTGGAGAAATGGAGGAGAGAGAAATCCAAAGT 1794  
Db |||||  
1621 TTCTAACCCACGTGAGAACCATGTGTTGGAGAAATGGAGGAGAGAGAAATCCAAAGT 1680  
Qy |||||  
1795 TCCTGATGATCTCATTTGAGCTCTCGATCCAGTCTTTCTGAGCTGTGTTCTCTGG 1854  
Db |||||  
1681 TCCTGATGATCTCATTTGAGCTCTCGATCCAGTCTTTCTGAGCTGTGTTCTCTGG 1740  
Qy |||||  
1855 ACTTTTTCATGATGTCAGGCAATAAATGCTTTTCATTTCTTTG 1896  
Db |||||  
1741 ACTTTTTCATGATGTCAGGCAATAAATGCTTTTCATTTCTTTG 1782

RESULT 3  
AX089606  
LOCUS AX089606 2873 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 6 from Patent WO0116314.  
ACCESSION AX089606  
VERSION AX089606.1 GI:13443798  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Wallach,D., Malinin,N., Sinha,I.W. and Leu,S.  
TITLE Iren protein, its preparation and use  
JOURNAL Patent: WO 0116314-A 6 08-MAR-2001;  
YEDA RESEARCH AND DEVELOPMENT CO. LTD. (IL)  
FEATURES  
source 1. .2873  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 81.6%; Score 1555; DB 6; Length 2873;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;  
Qy 118 ATGAGCGGATCNCNAGCAATGCAAAAGCAATTTCTGCTGGAGCGACTGCTGGATGCA 177  
|||||

Db 1 ATGAGCGATCACAGAAACAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60  
Qy 178 GTGAAAACAGTGCCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 237  
Db 61 GTGAAAACAGTGCCAGATCCGCTTTGGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 120  
Qy 238 AGGGTCACCTGCTGTGTGCGCCAGTTTGAAAGCGGTCTGTGACGATGGCTTGAAGAGAGT 297  
Db 121 AGGGTCACCTGCTGTGTGCGCCAGTTTGAAAGCGGTCTGTGACGATGGCTTGAAGAGAGT 180  
Qy 298 CGAGGATTTGGCACTCACAGCGGCAGCGATCAAGCAGGCGAGCGGGCTTTGGCAGCAAAACC 357  
Db 181 CGAGGATTTGGCACTCACAGCGGCAGCGATCAAGCAGGCGAGCGGGCTTTGGCAGCAAAACC 240  
Qy 358 GAAACAGAGCCGTGTTCTGTGTACTAAGAGAGAGTCTCAACAGACGACGACTGCAG 417  
Db 241 GAAACAGAGCCGTGTTCTGTGTACTAAGAGAGAGTCTCAACAGACGACGACTGCAG 300  
Qy 418 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGTGCGGCCCTGCTGCGC 477  
Db 301 CGCTTCTACTCCCTGCGCCACATCGCCTCAGACGTGGGCCGGGTGCGGCCCTGCTGCGC 360  
Qy 478 TGTGCCCTCAAGCAACTCCCTGGAGCGCTACCTGCAATGCTCTGCGCGACCGCTGC 537  
Db 361 TGTGCCCTCAAGCAACTCCCTGGAGCGCTACCTGCAATGCTCTGCGCGACCGCTGC 420  
Qy 538 AGGCTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGTCCAGTATG 597  
Db 421 AGGCTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGTCCAGTATG 480  
Qy 598 CTTTCTACATGGCAGCAGGTCTGAATCCATACCTCTTTTGGGATTTAAATCGACAAACAAG 657  
Db 481 CTTTCTACATGGCAGCAGGTCTGAATCCATACCTCTTTTGGGATTTAAATCGACAAACAAG 540  
Qy 658 GATTTGAACGGGCAGAGTAAGTTTCTCCACCGTTTCAGACCTTTAAAGGATCAACG 717  
Db 541 GATTTGAACGGGCAGAGTAAGTTTCTCCACCGTTTCAGACCTTTAAAGGATCAACG 600  
Qy 718 CAGAACGTGACCTCTCTGAGGAGTCCAGCAAGAGTGAGCAGCCTGTTCAGGGAG 777  
Db 601 CAGAACGTGACCTCTCTGAGGAGTCCAGCAAGAGTGAGCAGCCTGTTCAGGGAG 660  
Qy 778 ATCACAGCCTCTCTGCGCTCTCCATCTCATCAAACTGAAAGGACCGA-CCCTTG 836  
Db 661 ATCACAGCCTCTCTGCGCTCTCCATCTCATCAAACTGAAAGGACCGA-CCCTTG 720  
Qy 837 CTTGCTGTGTCAGGAATGTCAGTCTGATGCAATGCAAAAGGAGCGGAGAGAA 896  
Db 721 CTTGCTGTGTCAGGAATGTCAGTCTGATGCAAAAGGAGCGGAGAGAGAA 780  
Qy 897 AAGAAGTGACCAACAATATCTCATTTGATGATGAGGAAGTGAAGCAACTCTGGGAC 956  
Db 781 AAGAAGTGACCAACAATATCTCATTTGATGATGAGGAAGTGAAGCAACTCTGGGAC 840  
Qy 957 GTGTTTTAAAGACACTTGGGCGAGGGAGCTCAGAGGACAACTCCGACCGCTCTCT 1016  
Db 841 GTGTTTTAAAGACACTTGGGCGAGGGAGCTCAGAGGACAACTCCGACCGCTCTCT 900  
Qy 1017 GTCAATATCATGTCGCTTGAAGCCCTTTCGGGCTTACTCCAATGGAA-TCAGAGC 1075  
Db 901 GTCAATATCATGTCGCTTGAAGCCCTTTCGGGCTTACTCCAATGGAAAGTCAAGC 960  
Qy 1076 AGCAACTCATGAAAATTTGATTTCCCTGTCTTTTGAACGGGGAGTTTGGGTACCAAGAGCTT 1135  
Db 961 AGCAACTCATGAAAATTTGATTTCCCTGTCTTTTGAACGGGGAGTTTGGGTACCAAGAGCTT 1020  
Qy 1136 GATGTGAAAGCATTCGATGATGAAGATGGAATGAAAAGAGATGACGTGTATGAAAC 1195  
Db 1021 GATGTGAAAGCATTCGATGATGAAGATGGAATGAAAAGAGATGACGTGTATGAAAC 1080  
Qy 1196 TCATCAGGAAGGAGCAGAGGGGCCACTCGGAGTTCGCCGAGAGCACTGGAAGGGAC 1255  
Db 1081 TCATCAGGAAGGAGCAGAGGGGCCACTCGGAGTTCGCCGAGAGCACTGGAAGGGAC 1140

Qy 1256 ACCTGCTCTCCAGATGACAGCTGGGTCCGCTGAAAGGTGCTGCACATGACTCCGAC 1315  
Db 1141 ACCTGCTCTCCAGATGACAGCTGGGTCCGCTGAAAGGTGCTGCACATGACTCCGAC 1200  
Qy 1316 ATCTCTTCTCCCTGCTAGTGGCGTCTCTACAGCCCAAGCAGATGCCCTCCGGAAGC 1375  
Db 1201 ATCTCTTCTCCCTGCTAGTGGCGTCTCTACAGCCCAAGCAGATGCCCTCCGGAAGC 1260  
Qy 1376 CTGGAGAACGGGACAGGACAGGAGGACAGGTTCTCCGGATCCTCGGTACAGT 1435  
Db 1261 CTGGAGAACGGGACAGGACAGGAGGACAGGTTCTCCGGATCCTCGGTACAGT 1320  
Qy 1436 GTGGAAGCCAGCTCTCCAGGCCACGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1494  
Db 1321 GTGGAAGCCAGCTCTCCAGGCCACGAAGTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1380  
Qy 1495 GTGCCAGAGTCCATGACAAATTAGTAACTGCGCCAGGCCACTGTGGCATGATGAACAG 1554  
Db 1381 GTGCCAGAGTCCATGACAAATTAGTAACTGCGCCAGGCCACTGTGGCATGATGAACAG 1440  
Qy 1555 AAGGATGAGCTGGAGGAGGAGACAGATCACTCGGAAACCTGCTCGACGCTGAGATGGAG 1614  
Db 1441 AAGGATGAGCTGGAGGAGGAGACAGATCACTCGGAAACCTGCTCGACGCTGAGATGGAG 1500  
Qy 1615 CACTCAGCCCGCTCCGGCAAGAGGTGGACACCTTGAAGAGAGGTGGCTGAACAGGAG 1674  
Db 1501 CACTCAGCCCGCTCCGGCAAGAGGTGGACACCTTGAAGAGAGGTGGCTGAACAGGAG 1560  
Qy 1675 GAGCGGAGGCGATGAAGTCCAGGCGCTGGCCAG 1709  
Db 1561 GAGCGGAGGCGATGAAGTCCAGGCGCTGGCCAG 1595

RESULT 4  
AX089605  
LOCUS AX089605 3139 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 5 from Patent WO0116314.  
ACCESSION AX089605  
VERSION AX089605.1 GI:13443797  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Wallach, D., Malinin, N., Sinha, I.W. and Leu, S.  
TITLE Iren protein, its preparation and use  
JOURNAL YEDA: WO 0116314-A 5 08-MAR-2001;  
YEDA RESEARCH AND DEVELOPMENT CO. LTD. (IL)  
FEATURES  
Location/Qualifiers  
source 1..3139  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 81.6%; Score 1555; DB 6; Length 3139;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1588; Conservative 0; Mismatches 4; Indels 3; Gaps 3;  
Qy 118 ATGAGCGGATCNCNGAACNATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 177  
Db 1 ATGAGCGGATCACAAGAACAAATGACAAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCA 60  
Qy 178 GTGAAAACAGTCCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 237  
Db 61 GTGAAAACAGTCCAGATCCGCTTTNGAGGGAGAAAGAGATTGCTCGGATTCGACAGC 120  
Qy 238 AGGGTCACCTGCTGTGTGCGCCAGTTTGAAGCCGTCTGTGACGATGGCTTGAAGAGAGT 297  
Db 121 AGGGTCACCTGCTGTGTGCGCCAGTTTGAAGCCGTCTGTGACGATGGCTTGAAGAGAGT 180

QY 298 CGAGGATTGGCACTCAGCGGCGAGCGATCAAGCAGCGCGGGCTTTGCGCAGCAAAACC 357  
 Db 181 CGAGGATTGGCACTCAGCGGCGAGCGATCAAGCAGCGCGGGCTTTGCGCAGCAAAACC 240  
 QY 358 GAAACAGAGCCGCTGTTCTGTACTAGCTGAAGAGGTCTCTCAACAGACAGAGCTGCAG 417  
 Db 241 GAAACAGAGCCGCTGTTCTGTACTAGCTGAAGAGGTCTCTCAACAGACAGAGCTGCAG 300  
 QY 418 CGCTTTACTCTCCCTGCGCCACATCGCTTCAGACGCTGGCGCGGGTTCGCGCTTGGCTGCGC 477  
 Db 301 CGCTTTACTCTCCCTGCGCCACATCGCTTCAGACGCTGGCGCGGGTTCGCGCTTGGCTGCGC 360  
 QY 478 TGTGCGCTTCAACGAACTCCCTGGAGCGTACTCTGCACATGCTCTGCGCGGAGCCGCTGC 537  
 Db 361 TGTGCGCTTCAACGAACTCCCTGGAGCGTACTCTGCACATGCTCTGCGCGGAGCCGCTGC 420  
 QY 538 AGGCTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATG 597  
 Db 421 AGGCTGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATG 480  
 QY 598 CTTCTTACCATTGGCAGAGTCTGAATCCATCTCTTTTCCGATTAACATCGAACAAAG 657  
 Db 481 CTTCTTACCATTGGCAGAGTCTGAATCCATCTCTTTTCCGATTAACATCGAACAAAG 540  
 QY 658 GATTTGAACGGCGAGAGTAAGTTGCTCCACCGTTTCAGACCTCTTAAAGGAGTCAAG 717  
 Db 541 GATTTGAACGGCGAGAGTAAGTTGCTCCACCGTTTCAGACCTCTTAAAGGAGTCAAG 600  
 QY 718 CAGAACTGACCTCTTCTGTGAAGAGTCCACGCAAGAGTGAGCAGCTGTTTCAGGGAG 777  
 Db 601 CAGAACTGACCTCTTCTGTGAAGAGTCCACGCAAGAGTGAGCAGCTGTTTCAGGGAG 660  
 QY 778 ATCAGACCTCTCTGCGGCTCTCCATCTCTCAAACTGAAACAGAGACCGA-CCCTTG 836  
 Db 661 ATCAGACCTCTCTGCGGCTCTCCATCTCTCAAACTGAAACAGAGACCGA-CCCTTG 720  
 QY 837 CTTGCTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGAGGCGGAGAGAA 896  
 Db 721 CTTGCTGTCCAGGAATGTCAGTGTGATGCCAAATGCAAAAGAGGCGGAGAGAA 780  
 QY 897 AAGAAAGTGACCAATACTCTATTGATGATGAGGAAGATGAGCAGAGACTCTGGGGAC 956  
 Db 781 AAGAAAGTGACCAATACTCTATTGATGATGAGGAAGATGAGCAGAGACTCTGGGGAC 840  
 QY 957 GTGTTTAAAGACACCTGGGGCAGGGGAGAGCTCAGAGCAAACTCCGACCGCTCTCT 1016  
 Db 841 GTGTTTAAAGACACCTGGGGCAGGGGAGAGCTCAGAGCAAACTCCGACCGCTCTCT 900  
 QY 1017 GTCATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAACTCCCAATGAA-TCAGAGC 1075  
 Db 901 GTCATATCATGTCCGCTTTGAAAGCCCTTTCGGGCTTAACTCCCAATGAA-TCAGAGC 960  
 QY 1076 AGCAACTCATGGAATAATGATTCCTCTCTTTGACCGGGAGTTTGGGTACCAAGAGCTT 1135  
 Db 961 AGCAACTCATGGAATAATGATTCCTCTCTTTGACCGGGAGTTTGGGTACCAAGAGCTT 1020  
 QY 1136 GATGTGAAAGACATCGATGATGAAGATGTGATGAAACGAAGATGACGTGTATGGAAC 1195  
 Db 1021 GATGTGAAAGACATCGATGATGAAGATGTGATGAAACGAAGATGACGTGTATGGAAC 1080  
 QY 1196 TCATCAGGAAGGACACAGGGGCCATTCGAGTTCGCGCGCGAGAGCCACTGGAAGGAAAC 1255  
 Db 1081 TCATCAGGAAGGACACAGGGGCCATTCGAGTTCGCGCGCGAGAGCCACTGGAAGGAAAC 1140  
 QY 1256 ACCTGCTCTCCAGATGACAGTGGGCTCCGCTGAGCTGCTGCAATGATCTCGAC 1315  
 Db 1141 ACCTGCTCTCCAGATGACAGTGGGCTCCGCTGAGCTGCTGCAATGATCTCGAC 1200  
 QY 1316 ATCTCTTCTCTGTGAGTGGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAGC 1375  
 Db 1201 ATCTCTTCTCTGTGAGTGGGCTCTTACAGCCAGCAGATGCCCCCTCGGAAGC 1260  
 QY 1376 CTGGAGAAACGGGACAGGACCAAGTCTTCTCCCGATCTCTGCACTTCGGTACAGT 1435

Db 1261 CTGGAGAACGGGACAGGACAGGACCGAGTTCTCCGGATCTCGACTTCGGTACGT 1320  
 QY 1436 GTGGAAGCCAGCTCTCTCAGGCCACGGAAGTCTCTCTGAGCAGCCTGTTA-CTTCTGCTCA 1494  
 Db 1321 GTGGAAGCCAGCTCTCTCAGGCCACGGAAGTCTCTCTGAGCAGCCTGTTACCTTCTGCTCA 1380  
 QY 1495 GTGCCAGAGTCCATGACAAATTAGTAACTGCGGCAGGCCACTGTGCCATGATGAACAGG 1554  
 Db 1381 GTGCCAGAGTCCATGACAAATTAGTAACTGCGGCAGGCCACTGTGCCATGATGAACAGG 1440  
 QY 1555 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTGACAGCTGAGATGGAG 1614  
 Db 1441 AAGGATGAGCTGGAGGAGGAGAACAGATCACTCGGAAACCTGTGACAGCTGAGATGGAG 1500  
 QY 1615 CACTCAGCCGCTCTCCGCAAGAGGTGGACACCTTGAAGAGGAGGTGGCTGAACAGGAG 1674  
 Db 1501 CACTCAGCCGCTCTCCGCAAGAGGTGGACACCTTGAAGAGGAGGTGGCTGAACAGGAG 1560  
 QY 1675 GAGCGGAGGCGATGAAGGTTCAGGCGCTGGCCAG 1709  
 Db 1561 GAGCGGAGGCGATGAAGGTTCAGGCGCTGGCCAG 1595

RESULT 5  
 BD157657  
 LOCUS  
 DEFINITION 2248 bp DNA linear PAT 17-JAN-2003  
 Primer for synthesizing full-length cDNA and use thereof.  
 ACCESSION  
 BD157657  
 VERSION  
 BD157657.1 GI:27863415  
 KEYWORDS  
 JP 2002191363-A/12500.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2248)  
 Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,  
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.  
 Primer for synthesizing full-length cDNA and use thereof  
 Patent: JP 2002191363-A 12500 09-JUL-2002;  
 HELIX RESEARCH INSTITUTE  
 OS Homo sapiens (human)  
 FN JP 2002191363-A/12500  
 PD 09-JUL-2002  
 PF 28-JUL-2000 JP 2000280990  
 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU  
 SAITO,  
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,  
 KEIICHI NAGAI, TETSUJI OTSUKI  
 PC  
 C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC  
 10,  
 C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/09 CC  
 Primer for synthesizing full-length cDNA and use thereof PH Key

FT CDS Location/Qualifiers  
 (63)..(1187).  
 1..2248  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

FEATURES  
 source  
 1..2248  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 59.7%; Score 1138.2; DB 6; Length 2248;  
 Best Local Similarity 94.7%; Pred. No. 5.4e-271;  
 Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;  
 QY 77 GCGCGCGCGCGCGCAGCAGCCCGCCCGGGAGAGCCATGAGCGGATCNGAACN 136  
 Db 22 GCGCGCGCGCGCGCAGCAGCCCGCCCGGGAGAGCACCATGAGCGGATCAGAGACA 81  
 QY 137 ATGACAAAAGACAAATTTCTGCTGAGGAGTCTGATGATGAGTGAACAGTGCAGATCC 196

```
Db      82  ATGCAAAAAGACAAATTTCTGTGGAGCGACTGCTGGATGCAAGTGAACAGTGCAGATCC 141
Qy      197  GCTTTNGAGGGAAGAGATTGCTCGGATTCGACAGAGGTCACCTGTCTGTGTG 256
Db      142  GCTTTGGAGGGAAGAGATTGCTCGGATTCGACAGAGGTCACCTGTCTGTGTG 201
Qy      257  CCCAGTTTGAAGCCGCTCTGAGCATGGCTTGAAGAGGATCGAGATTGGCACTCACAG 316
Db      202  CCCAGTTTGAAGCCGCTCTGAGCATGGCTTGAAGAGGATCGAGATTGGCACTCACAG 261
Qy      317  CGGAGCGATCAAGCAGGAGCGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 376
Db      262  CGGAGCGATCAAGCAGGAGCGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 321
Qy      377  GGTACTAGTGAAGAGTCTCTACACAGCAGCAGCTCAGGCTTCTACTCCCTGCGCC 436
Db      322  GGTACTAGTGAAGAGTCTCTCAACAGCAGCAGCTCAGGCTTCTACTCCCTGCGCC 381
Qy      437  ACATCGCTCAGACGCTGGGCGGGGTCGGCTGGCTGGCTGGCTCAACGAACACT 496
Db      382  ACATCGCTCAGACGCTGGGCGGGGTCGGCTGGCTGGCTGGCTCAACGAACACT 441
Qy      497  CCTGGAGGCTACCTGACATGCTCTCGGCGACCGCTGAGGCTGAGCACTTTTATG 556
Db      442  CCTGGAGGCTACCTGACATGCTCTCGGCGACCGCTGAGGCTGAGCACTTTTATG 501
Qy      557  AAGACTGGTCTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATTGCGAG 616
Db      502  AAGACTGGTCTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATTGCGAG 561
Qy      617  GTCTGAATCCACTACTCTTTGCGATTACATCGACCAACAGGATTTGAACGGGAGAGTA 676
Db      562  GTCTGAATCCACTACTCTTTGCGATTACATCGACCAACAGGATTTGACGGGAGAGTA 621
Qy      677  AGTTTGCTCCACCGTTTTCAGACTCTTAAAGAGTCAACGAGAACTGACCTCCTTGC 736
Db      622  AGTTTGCTCCACCGTTTTCAGACTCTTAAAGAGTCAACGAGAACTGACCTCCTTGC 681
Qy      737  TGAAGGAGTCCACGAGGATGAGCAGCTGTTTCAAGGAGTCAACGAGAACTGACCTCCTGCG 796
Db      682  TGAAGGAGTCCACGAGGATGAGCAGCTGTTTCAAGGAGTCAACGAGAACTGACCTCCTGCG 741
Qy      797  TCTCATCTCTCATCAAACTGAACAGGAGACCGA - CCTTTCCTGTGTGTGTCAGGAATG 855
Db      742  TCTCATCTCTCATCAAACTGAACAGGAGACCGA - CCTTTCCTGTGTGTGTCAGGAATG 801
Qy      856  TCAGTGTCTGATGCCAAATGCAAAAGGAGCGGAGAGAGAAAGAGTGACCAACATAA 915
Db      802  TCAGTGTCTGATGCCAAATGCAAAAGGAGCGGAGAGAGAGAAAGAGTGACCAACATAA 861
Qy      916  TCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGTTTAAAGACACCTG 975
Db      862  TCTCATTTGATGAGGAAGATGAGCAGAACTCTGGGAGCTGTTTAAAGACACCTG 921
Qy      976  GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 1035
Db      922  GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 981
Qy      1036  TTGAAGCCCTTCGGGCTTAACCTCAATGAA - TCAGAGCAGCACTCATGTGAAATTTG 1094
Db      982  TTGAAGCCCTTCGGGCTTAACCTCAATGAAAGTCAAGAGTCAAGCACTCATGTGAAATTTG 1041
Qy      1095  ATTCCCTGTCTTTGAACGGGAGTTTGGGTACCAAGAGCTTGATGTGAAGAGCATCATG 1154
Db      1042  ATTCCCTGTCTTTGAACGGGAGTTTGGGTACCAAGAGCTTGATGTGAAGAGCATCATG 1101
Qy      1155  ATGAAGATGTGGATCAAAACGAAGATGACGTGTGTGAAACTCATCAGGAAGGAGCACA 1214
Db      1102  ATGAAGATGTGGATCAAAACGAAGATGACGTGTGTGAAACTCATCAGGAAGGAGCACA 1161
Qy      1215  GGGGCCACTCGGAGTCGCGGAGAGCACTGGAAGGGAACCACTGCTTCCAGATGC 1274
Db      1162  GGGGCCACTCGGAGTCGCGGAGAGTAAAGTTTGTGTGTAAGGTGGAGTCTCACCCTCC 1221
```

```
Qy      1275  ACAGCTGGCTCCGCTGAAGGTGCTGCACATGATGATCGGACATCTTCTCCCTGTCAGTG 1334
Db      1222  CCAGGCTGGAGTGCCGTGGCGTGATCTCAGCTCACTGCAACTCCACCTCCGGTTTCAG 1281
Qy      1335  GC 1336
Db      1282  GC 1283

RESULT 6
LOCUS      AX879111                      2248 bp      DNA      linear      PAT 17-DEC-2003
DEFINITION Sequence 14016 from Patent EP1074617.
ACCESSION  AX879111
VERSION     AX879111.1      GI:40033847
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
            Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
  TITLE     Primers for synthesising full-length cDNA and their use
  JOURNAL   Patent: EP 1074617-A 14016 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES
  source    1..2248
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            63..1190
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAE930370.1"
            /db_xref="GI:40033848"
            /translation="MSGSONNDKQELLELLDAVKOCQIRFGGRKEIASDSRVTC
            LCAQFVAVLQHLKRSGLATAAAIKQAAGFAKKTETEPVFVYKEVLLKELQRF
            YSRHLASDVGRGRAMRLCALNEASUYELHMLADRCLSTFYFVWDEBSLRF
            LPTAAGLSILIFAINIDNKLDQSKFAPVSDLLKSTQNTVSLKSTQGVSSIF
            RETTASAVSILIKPQETDPLPVSRNSADAKKERRKKKVTNIIIFDDEEDQ
            NSGVFKTTPGAGESSEDNSRNSVIMSAFSPFGPNSNGSSNSWKIDSLSLNGE
            FGQKLDVKIGDDEVDENEDDYVYGNSSGRKRGHSSEPK"

CDS
  1..2248
  /note="unnamed protein product"
  /codon_start=1
  /protein_id="CAE930370.1"
  /db_xref="GI:40033848"
  63..1190
  /note="unnamed protein product"
  /codon_start=1
  /protein_id="CAE930370.1"
  /db_xref="GI:40033848"

Query Match      59.7%; Score 1138.2; DB 6; Length 2248;
Best Local Similarity 94.7%; Pred. No. 5.4e-271;
Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;

Qy      77  GCCGGCGCGCGCGCANGCACCGGCCCGGGANAGCNCATGAGCGGATCNGAACN 136
Db      22  GCGGCGCGCGCGCGCANGCACCGGCCCGGGAGAGCACCATGAGCGGATCAGAGAAC 81
Qy      137  ATGACAAAAGACAAATTTCTGTGAGCGCACTGCTGGATGAGTGAAGAGTGCAGATCC 196
Db      82  ATGACAAAAGACAAATTTCTGTGAGCGCACTGCTGGATGAGTGAAGAGTGCAGATCC 141
Qy      197  GCTTTNAGGGAAGAGGAGATTGCTTCGATTCCGATCCGACAGAGGTCACTGTCTGTGTG 256
Db      142  GCTTTGAGGAGAGAGGAGATTGCTTCGATTCCGATCCGACAGAGGTCACTGTCTGTGTG 201
Qy      257  CCAGATTGAAGCGCTCTCGACATGGCTTGAGAGAGTTCGAGGATTCGACTTCACAG 316
Db      202  CCAGATTGAAGCGCTCTCGACATGGCTTGAGAGAGTTCGAGGATTCGACTTCACAG 261
Qy      317  CGGAGCGATCAAGCAGGAGCGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 376
Db      262  CGGAGCGATCAAGCAGGAGCGGCTTTGCCAGCAAAACCGAAACAGAGCCCGTGTCT 321
Qy      377  GGTACTAGTGAAGGAGTCTCTCAACAGCAGGATGCGAGCGCTTCTACTTCCCTGCGCC 436
Db      322  GGTACTAGTGAAGGAGTCTCTCAACAGCAGGATGCGAGCGCTTCTACTTCCCTGCGCC 381
```





/protein\_id="BAB14033.1"  
/db\_xref="GI:10433819"  
/translation="MSGSONDKRQFLRLERLLDAVKQCIIRFGGRKEIASDSRVTC  
LCAQFEAVLHGRLKSRGLALATAAIAKQAGPASKTETEPVPMVYKVLNKHLELRF  
YSLRHASDVGRGRLWALNEHSLERYLHMLLADRCRLSTFTFEDWSFVMDERSSM  
LPTMAAGLNSILFPAINDKOLDGQSKFAPTVSDLLKESTQNTVSLKESQGVSSLF  
REITASAVSILIKPEQETDPLPVVSRNVSDAKCKKRRKKKVTNIISFDDDEDEQ  
NSGDFKTPGAGESSESDRSSVNIIMAFESFPFNPNSGSSQSNKIDSLSLNGE  
FGYQKLDVKGIDDEDVEDNEDDVYNGSSGRKHRGHSPEK"

ORIGIN

Query Match 59.7%; Score 1138.2; DB 9; Length 2248;  
Best Local Similarity 94.7%; Pred. No. 5.4e-271;  
Matches 1195; Conservative 0; Mismatches 65; Indels 2; Gaps 2;  
  
QY 77 GCGCGCGCGCGCGCANGCACCGCCGCGGGANAGCNCCATGAGCGGATCNGNACN 136  
DB 22 GCGCGCGCGCGCGCAGCACCGCCCGGGGAGGACCACATGAGCGGATCAGAAACA 81  
  
QY 137 ATGACAAAGACAATTTCTGTGGAGCGACTGCTGGATGCAATGAAACAGTGCAGATCC 196  
DB 82 ATGACAAAGACAATTTCTGTGGAGCGACTGCTGGATGCAATGAAACAGTGCAGATCC 141  
  
QY 197 GCTTTNGAGGAGAAAGAGATTGCTCGGATTCGACAGCAGGAGTCACTGTCTGTGTG 256  
DB 142 GCTTTGGAGGAGAAAGAGATTGCTCGGATTCGACAGCAGGAGTCACTGTCTGTGTG 201  
  
QY 257 CCCAGTTTGAAGCCGCTCTGACGATGCTTGAAGGAGGTCGAGGATTTGCACATCACAG 316  
DB 202 CCCAGTTTGAAGCCGCTCTGACGATGCTTGAAGGAGGTCGAGGATTTGCACATCACAG 261  
  
QY 317 CGCGAGCGATCAAGCAGCAGCGGCTTTGCGCAAAACCGCAACAGAGCCCGTGTCT 376  
DB 262 CGCGAGCGATCAAGCAGCAGCGGCTTTGCGCAAAACCGCAACAGAGCCCGTGTCT 321  
  
QY 377 GGTACTAGCTGAGAGAGTCTCTCAACAGCAGGAGTGCAGCGCTTCTACTCCCTGCGCC 436  
DB 322 GGTACTAGTGAAGAGTCTCTCAACAGCAGGAGTGCAGCGCTTCTACTCCCTGCGCC 381  
  
QY 437 ACATCGCTCAGACGTGGCGCGGCTCGCGCTGCTGCGCTGCTCAACGAACACT 496  
DB 382 ACATCGCTCAGACGTGGCGCGGCTCGCGCTGCTGCGCTGCTCAACGAACACT 441  
  
QY 497 CCCTGGAGCGTACTGACATGCTCTGCGCGGACCGCTGCGAGCTGAGCACTTTTATG 556  
DB 442 CCCTGGAGCGTACTGACATGCTCTGCGCGGACCGCTGCGAGCTGAGCACTTTTATG 501  
  
QY 557 AAGACTGTCTTTGTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGCGACGAG 616  
DB 502 AAGACTGTCTTTGTGTGATGATGAAGAAAGTCCAGTATGCTTCTTACCATGCGACGAG 561  
  
QY 617 GTCTGAATCCATACTCTTTCCGATTAAATCGACAAAGGATTTGAAACGGGCGAGTA 676  
DB 562 GTCTGAATCCATACTCTTTCCGATTAAATCGACAAAGGATTTGAAACGGGCGAGTA 621  
  
QY 677 AGTTTGCTCCACCGTTTCAGACCTCTTAAAGAGTCAACCGCAGAACGTGACCTCTTGC 736  
DB 622 AGTTTGCTCCACCGTTTCAGACCTCTTAAAGAGTCAACCGCAGAACGTGACCTCTTGC 681  
  
QY 737 TGAAGGAGTCCACCAAGGAGTGAAGCGCTGTTTCAAGGAGATCAGACGCTCTCTGCGG 796  
DB 682 TGAAGGAGTCCACCAAGGAGTGAAGCGCTGTTTCAAGGAGATCAGACGCTCTCTGCGG 741  
  
QY 797 TCTCCATCTCTCATCAAACTGAAACAGGAGACCGA - CCCTTGCCTGTGCTGCTCCAGGAATG 855  
DB 742 TCTCCATCTCTCATCAAACTGAAACAGGAGACCGACCCCTTGCCTGTGCTGCTCCAGGAATG 801  
  
QY 856 TCAGTGTGATGTCCAATGCAAAAGGAGCGGAAGAAAGAAAGAAAGTCAACCAATAA 915  
DB 802 TCAGTGTGATGTCCAATGCAAAAGGAGCGGAAGAAAGAAAGTCAACCAATAA 861  
  
QY 916 TCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGACGTGTTTTAAAAAGACACTG 975

DB 862 TCTCATTTGATGATGAGGAGATGAGCAGAACTCTGGGAGCGTGTTTAAAAAGACACTG 921  
QY 976 GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 1035  
DB 922 GGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCAATATCATGTCCGCT 981  
QY 1036 TTGAAAGCCCTTCGCGCTTAACTCCAAATGGAA - TCAGAGCAGCAACTCATGGAATAATG 1094  
DB 982 TTGAAAGCCCTTCGCGCTTAACTCCAAATGGAACTCAGAGCAGCACTCATGGAATAATG 1041  
QY 1095 ATTCCCTGTCTTTGAACGGGAGTTTCGGTACCAGAACTTGATGTGAAAAGCATCGATG 1154  
DB 1042 ATTCCCTGTCTTTGAACGGGAGTTTCGGTACCAGAACTTGATGTGAAAAGCATCGATG 1101  
QY 1155 ATGAAGATGTGGATGAAAACGAAAGATGACCTGTATGAAAACATCATCAGGAAGGAGACA 1214  
DB 1102 ATGAAGATGTGGATGAAAACGAAAGATGACCTGTATGAAAACATCATCAGGAAGGAGACA 1161  
QY 1215 GGGGCCACTCGGAGTCCCGGAGAACCACTGGAAGGGAAACACCTGCTCTCCAGATGC 1274  
DB 1162 GGGGCCACTCGGAGTCCCGGAGAACCACTGGAAGGGAGTGTGTGTGAAGGTGGAGTCTCACCGTCCC 1221  
QY 1275 ACAGCTGGCTCCGCTGAAGGTGCTGCACAATGACTCCGACATCTCTTCCCTGTGATG 1334  
DB 1222 CCAGGCTGGAGTCCGCTGGGCTGATCTCAGTCACTGCAACCTCCACCTCCCGGTTTCCAG 1281  
QY 1335 GC 1336  
DB 1282 GC 1283  
  
RESULT 8  
CO714292  
LOCUS CO714292 1033 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 226 from Patent WO02068579.  
ACCESSION CO714292  
VERSION CO714292.1 GI:42275149  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
  
REFERENCE 1  
AUTHORS Venter, C. J., Adams, M. C., Li, P. W. and Myers, E. W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
human exons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 226 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source Location/Qualifiers  
1 . . 1033  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
  
ORIGIN  
Query Match 47.7%; Score 909.4; DB 6; Length 1033;  
Best Local Similarity 91.9%; Pred. No. 2.9e-214;  
Matches 1026; Conservative 0; Mismatches 5; Indels 86; Gaps 3;  
  
QY 125 GATCNCNAGCAATGACAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCAAGTCA 184  
DB 1 GATCAGACAAATGACAAAGACAATTTCTGCTGGAGCGACTGCTGGATGCAAGTCA 60  
QY 185 AGTGCAGATCCGCTTTNGAGGAGAAAGAGATTCCTCGGATTCGACACGAGGTCA 244  
DB 61 AGTGCAGATCCGCTTTNGAGGAGAAAGAGATTCCTCGGATTCGACACGAGGTCA 120  
QY 245 CTTGTCTGTGCCCCAGTTTGAAGCGCTCTGACGATGGCTTGAAGAGAGTCCAGGAT 304  
DB 121 CTTGTCTGTGCCCCAGTTTGAAGCGCTCTGACGATGCTTGAAGAGAGTCCAGGAT 180  
QY 305 TGGCACTCAGCGGCGAGCGATCAGCAGGACGGGCTTTTGCACGAAAAACAG 364



```
Db 181 TGGCACTCAGACGGCAGCATCAAGCAGGAGCGGGCTTGGCAGCAAAACCGAA--- 237
Qy 365 AGCCCGTGTCTGGTACTAGTGAAGAGGTCTCAACAGCAGCAGCTGCGGCTTCT 424
Db 238 ----- 237
Qy 425 ACTCCCTGCCACATCGCTCAGACGTGGGCGGGGTGCGGCTGCGTGTGCCC 484
Db 238 ----- 237
Qy 485 TCAACGAACACTCCCTGGAGCTTACTCTGACATGCTCCCTGGGCGGCGCTGAGGCTGA 544
Db 277 TCAACGAACACTCCCTGGAGCTTACTCTGACATGCTCCCTGGGCGGCGCTGAGGCTGA 336
Qy 545 GCACCTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTA 604
Db 337 GCACCTTTTATGAAGACTGTCTTTTGTGATGATGAAGAAAGTCCAGTATGCTTCTTA 396
Qy 605 CCATGGCAGAGTCTGAACCTCATACTCTTTTGGATTAATACATCGACAAACAGATTGA 664
Db 397 CCATGGCAGAGTCTGAACCTCATACTCTTTTGGATTAATACATCGACAAACAGATTGA 456
Qy 665 ACCGGCAGAGTAAAGTTTGTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAACG 724
Db 457 ACCGGCAGAGTAAAGTTTGTCCACCGTTTCAGACCTCTTAAAGGAGTCAACGCAACG 516
Qy 725 TGACCTCTTGTCTGAAGAGTCCACGCAAGAGTGAAGCCTGTTTCAGGAGATCACAG 784
Db 517 TGACCTCTTGTGAAGAGTCCACGCAAGAGTGAAGCCTGTTTCAGGAGATCACAG 576
Qy 785 CTTCTCTGCGCTCTCATCTCATCAAACTGAACAGGAGACCGA---CCCTTCCCTGTGC 843
Db 577 CTTCTCTGCGCTCTCATCTCATCAAACTGAACAGGAGACCGA---CCCTTCCCTGTGC 636
Qy 844 TGTCCAGGAATGTCAGTGTCTGATGCTCCAAATGCAAAAGGAGCGGAAGAAAGAAAG 903
Db 637 TGTCCAGGAATGTCAGTGTCTGATGCTCCAAATGCAAAAGGAGCGGAAGAAAGAAAG 696
Qy 904 TGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGTTTA 963
Db 697 TGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGTTTA 756
Qy 964 AAAGACACTCGGGCAGGGAGAGTCTCAGAGCAAACTCCGACCGCTCTCTGTCAATA 1023
Db 757 AAAGACACTCGGGCAGGGAGAGTCTCAGAGCAAACTCCGACCGCTCTCTGTCAATA 816
Qy 1024 TCATGTCCGCTTTGAAGCCCTTTCGGGCTTAATCCAAATGGAA---TCAGAGCAGCACT 1082
Db 817 TCATGTCCGCTTTGAAGCCCTTTCGGGCTTAATCCAAATGGAA---TCAGAGCAGCACT 876
Qy 1083 CATGGAAAATTTGATTCCTCTGTTTGAACGGGAGTTTGGGTACCAAGCTTTGATGTGA 1142
Db 877 CATGGAAAATTTGATTCCTCTGTTTGAACGGGAGTTTGGGTACCAAGCTTTGATGTGA 936
Qy 1143 AAAGCATTCGATGATGAAGATGTGGATGAACAGGAGTACGTTGTAAGAACTCATCAG 1202
Db 937 AAAGCATTCGATGATGAAGATGTGGATGAACAGGAGTACGTTGTAAGAACTCATCAG 996
Qy 1203 GAAGGAAGCAGAGGGCCACTCGAGTTCGCCAGAA 1239
Db 997 GAAGGAAGCAGAGGGCCACTCGAGTTCGCCAGAA 1033
```

## RESULT 9

```
HS0808428
LOCUS      5833 bp      mRNA      linear      PRI 30-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686E09125 (from clone DKFZp686E09125).
ACCESSION  BX648280
VERSION     BX648280.1  GI:34367439
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
```

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 5833)  
Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.  
The German Human cDNA Consortium  
Direct Submission  
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.  
This clone (DKFZp686E09125) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

## FEATURES

Location/Qualifiers  
source  
1..5833  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686E09125"  
/tissue\_type="human uterus"  
/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_sfi; host DH10B; sites SfiIIA + SfiIB"  
/dev\_stage="adult"  
polya\_signal  
5789..5794  
polya\_site  
5810

## ORIGIN

Query Match 41.7%; Score 794.4; DB 9; Length 5833;  
Best Local Similarity 96.9%; Pred. No. 1e-185;  
Matches 853; Conservative 0; Mismatches 21; Indels 6; Gaps 4;  
QY 363 AGAGCCCGTGTCTTGTGTTACTAGTGAAGGAGTCTCTCAACAGCAGCAGTGC-AGCGCT 421  
Db 1381 AGAGCCCGTGTCTTGTGTTACTAGTGAAGGAGTCTCTCAACAGCAGCAGTGC 1440  
QY 422 TCTACTCCCTCGGCCACATCGCTCAGAGTGGGCGGGGTGCGGCTGGCTGGCTGTG 481  
Db 1441 TCTACTCCCTCGGCCACATTTGCTCAGAGTGGGCGGGGCGGCACCTGGCTGCTGTG 1500  
QY 482 CCCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGTCTCTGGCGGCGGCTGCAGGC 541  
Db 1501 CCCTCAACGAACACTCCCTGGAGCGCTACTCTGCACATGTCTCTGGCGGCGGCTGCAGGC 1560  
QY 542 TGAGCACTTTTATGAAGACTGTCTTTTGTGATGGATGAAGAAAGTCCAGTATGCTTC 601  
Db 1561 TCAGTACTTTTATGAAGACTGTCTTTTGTGATGGATGAAGAGAGGTCCAGTATGCTTC 1620  
QY 602 CTACCATGGCAGCAGTCTGAACCTCATACTCTTTGGGATTAACATCGACAAACAGATT 661  
Db 1621 CTACCATGGCAGCAGTCTCGAACTCCATACTCTTTTGGGATTAACATCGACAAACAGATT 1680  
QY 662 TGAACGGGCGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTTAAAGGAGTCAACGAGA 721  
Db 1681 TGAACGGGCGAGAGTAAAGTTTGTCTCCACCGTTTTCAGACCTCTTTAAAGGAGTCAACGAGA 1740  
QY 722 AGGTGACCTCTTTGTGTTGAAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGAGATCA 781  
Db 1741 ATGTGA---CCTTGTCTGAAGGAGTCCACGCAAGGAGTGAAGCAGCTGTTTCAGGAGATCA 1797  
QY 782 CAGCTCTCTCTGCGCTCTCATCTCATCAACCTGAAACAGGAGACCGA---CCCTTGGCTG 840  
Db 1798 CAGCTCTCTCTGCGCTCTCATCTCATCAACCTGAAACAGGAGACCGACCCCTTGGCCG 1857  
QY 841 TCGTGTCCAGGAGTGTCACTGTGTGATGCCAAATGCAAAAGGAGGAGCGGAAGAAAGAA 900  
Db 1858 TCGTGTCCAGGAGTGTCACTGTGTGATGCCAAATGCAAAAGGAGGAGCGGAAGAAAGAA 1917



FEATURES		Location/Qualifiers	
source		1..6045	/organism="Homo sapiens"
			/mol_type="unassigned DNA"
			/db_xref="taxon:9606"
CDS		364..957	
			/note="unnamed protein product"
			/codon_start=1
			/protein_id="CAE91902.1"
			/db_xref="GI:40038171"
			/translations="MRVHCAIIPSDMLHISTNCRTAEBMVLTLDDYLFHFEVQAVSSL SGQKHGKQOLDPLTYIGIRCHLFYKFGFISDMYRIKOSIDSKCHTARRKQWGSL VYKFSRRPNSSYCPSPMSTPPARELLQPOLQPOLHVALANVOVQVLIHQICE DGOVQVPIQGLHIAKVPQEQVQITQDSSEPCSGTT"
ORIGIN			
Query Match		41.6%; Score 792.8; DB 6; Length 6045;	
Best Local Similarity		96.8%; Pred. No. 2.6e-185;	
Matches 852; Conservative		0; Mismatches 22; Indels 6; Gaps 4;	
QY	363	AGAGCCCGTGTTCCTGCTACGTGAAGGAGGTCTCAACAAGCAGCAGCTGC-AGCGCT	421
Db	932	AAAGCCCGTGTTCCTGCTACGTGAAGGAGGTCTCAACAAGCAGCAGCTGC	991
QY	422	TCTACTCCCTCGCCACATCGCTCAGACGTGGCGCGGGTCCGCTCGCTGCTGTG	481
Db	992	TCTACTCCCTCGCCACATCGCTCAGACGTGGCGCGGGTCCGCTCGCTGCTGTG	1051
QY	482	CCCTCAACGAACACTCCCTGGAGCGTACTCTGCACATGCTCTCGGCGGACCGCTG	541
Db	1052	CCCTCAACGAACACTCCCTGGAGCGTACTCTGCACATGCTCTCGGCGGACCGCTG	1111
QY	542	TGAGCACTTTTATGAAGACTGCTCTTTGTGTGATGATGAAGAAAGTCCAGTATGCTTC	601
Db	1112	TCAGTACTTTTATGAAGACTGCTCTTTGTGTGATGATGAAGAGTCCAGTATGCTTC	1171
QY	602	CTACCATGGCAGAGTCTGAATCTCAATCTTTTGGATTAATCATCGACAACAGGATT	661
Db	1172	CTACCATGGCAGAGTCTGAATCTCAATCTTTTGGATTAATCATCGACAACAGGATT	1231
QY	662	TGAACGGGCGAGTAAGTTTGCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACCGA	721
Db	1232	TGAACGGGCGAGTAAGTTTGCTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACCGA	1291
QY	722	ACGTGACCTCTTCTGCTGAAGAGTCCACGCAAGAGTGAGCAGCTGTTTCAGGAGATCA	781
Db	1292	ATGTGA---CCTTGCTGAAGAGTCCACGCAAGAGTGAGCAGCTGTTTCAGGAGATCA	1348
QY	782	CAGCCTCTCTGCGCTCCATCTCTCATCAAACTGAACAGGAGACCGA-CCTTTCCTTG	840
Db	1349	CAGCCTCTCTGCGCTCCATCTCTCATCAAACTGAACAGGAGACCGACCCCTTCCCG	1408
QY	841	TCGTGTCCAGGATGTCAGTGTGTGTCGCAATGCAAAAAGGAGCGGAAGAAAAAGA	900
Db	1409	TCGTGTCCAGGATGTCAGTGTGTGTCGCAATGCAAAAAGGAGCGGAAGAAAAAGC	1468
QY	901	AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCTGT	960
Db	1469	AAGTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGAGCATGT	1528
QY	961	TAAAAAGACACCTGGGGCAGGGAGAGTCTCAGAGGACAACTCCGACCGCTCTCTGTCA	1020
Db	1529	TAAAAAGACACCTGGGGCAGGGAGAGTCTCAGAGGACAACTCCGACCGCTCTCTGTCA	1588
QY	1021	ATATCATGTCCGCTTTGAAGCCCTTGGGGCTTAATCTCAATGAA-TCAGAGCAGCA	1079
Db	1589	ATATCATGTCCGCTTTGAAGCCCTTGGGGCTTAATCTCAATGAA-TCAGAGCAGCA	1648
QY	1080	ACTCATGGAATAATTGATTCCTCTCTTTGAACGGGAGTTTGGTACCAGAGCTTGATG	1139
Db	1649	ACTCATGGAATAATTGATTCCTCTCTTTGAACGGGAGTTTGGTACCAGAGCTTGATG	1708
QY	1140	TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAACATCAT	1199

Db	1709	TGAAAAGCATCGATGATGAAGATGTGGATGAAAACGAAGATGACGTGTATGGAACATCAT	1768
Qy	1200	CAGGAAGGAAGCAGCAGGGGCCACTCGAGTCGCCCGAGAA	1239
Db	1769	CAGGACGGAAGCAGCAGGGGCCACTCAGATCGCCCGAGAA	1808
RESULT 12			
AK023827			
LOCUS	Homo sapiens cDNA FLJ13765 fis, clone PLACE4000128, weakly similar		
DEFINITION	to Mus musculus putative transcription factor mRNA.		
ACCESSION	AK023827		
VERSION	AK023827.1 GI:10435879		
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Hirakawa, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Houta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terahashina, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs		
JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)		
PUBLISHED	14702039		
REFERENCE	2		
AUTHORS	Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Oshima, A. NEDO human cDNA sequencing project		
TITLE	Unpublished		
JOURNAL	3 (bases 1 to 6045)		
REFERENCE	Isogai, T. and Otsuki, T.		
AUTHORS	Direct Submission		
TITLE	Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)		
JOURNAL	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center		
COMMENT			

etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES	source	Location/Qualifiers
		1..6045
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="PLACE4000128"
		/tissue_type="placenta"
		/clone_lib="PLACE4"
		/note="cloning vector: pME18SFL3"
CDS		364..957
		/note="unnamed protein product"
		/codon_start=1
		/protein_id="BAB14692.1"
		/db_xref="GI:10435880"
		/translation="MRVHCALIIPSDMLHISTNCRTAEBEVLITLLDYLFHHEVQAVSSL SQGKHKKOLDPLTIVYRIFGCHLIFYKGTESDWYRKQSIDSKCTAWRRKQWGSL VWKSFRMNFSSSYCESEPMNSTPPARELLQPLQPOQALHYALANVQVQVHQHGE DGVQVYPOGHLLHIAKVPGQEYQVITQDSBSPCSGTT"
ORIGIN		
	Query Match	41.6%; Score 792.8; DB 9; Length 6045;
	Best Local Similarity	96.8%; Pred. No. 2.6e-185;
	Matches 852; Conservative	0; Mismatches 22; Indels 6; Gaps 4;
Qy	363	AGAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCTCTCAACAGCAGCAGCTGC-ACGCGT 421
Db	932	AAAGCCCGTGTTCTGGTACTACGTGAAGGAGGTCTCTCAACAGCAGCAGCTGCAAGCGCT 991
Qy	422	TCTACTCCCTGGCGCCACATCGCTCTCAGACGTGGCGGGGTCTCGGCTTGCCTGCGCTGTG 481
Db	992	TCTACTCCCTGGCGCCACATCGCTCTCAGACGTGGCGGGGGCGCACCTGGCTCGCTGTG 1051
Qy	482	CCCTCAACGAACTCCCTCGAGCGGTACTCTGCACATGCTCTCTGGCGGACCGCTGAGGC 541
Db	1052	CCCTCAACGAACTCCCTCGAGCGGTACTCTGCACATGCTCTCTGGCGGACCGCTGAGGC 1111
Qy	542	TGAGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCTCAGTATGCTTC 601
Db	1112	TCAGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGAGTCCAGTATGCTTC 1171
Qy	602	CTACCATGGCAGCAGGTCTGAATCCATATCTCTTTGGGATTAACATCGACAACAAGGATT 661
Db	1172	CTACCATGGCAGCAGGTCCGAATCCATATCTCTTTGGGATTAACATGTGACAACAAGGATT 1231
Qy	662	TGAACGGGCAGAGTAAGTTTGCTCCCAACCGTTTTCAGACCTCTTAAAGGAGTCAACGCAGA 721
Db	1232	TGAACGGGCAGAGTAAGTTTGCTCCCAACCGTTTTCAGACCTCTTAAAGGAGTCAACGCAGA 1291
Qy	722	ACGTGACCTCCTTGTGAAGGATCCACGCAAGGATGAGCAGCCTGTTTCAGGGAGATCA 781
Db	1292	ATGTGA---CCTTGTCTGAAGGATCCACGCAAGGATGAGCAGCCTGTTTCAGGGAGATCA 1348
Qy	782	CAGCCTCTCTGCGCTTCCATCTCATCAAACTGAACAGGAGACCGA-CCCTTGCCCTG 840
Db	1349	CAGCCTCTCTGCGCTTCCATCTCATCAAACTGAACAGGAGACCGACCCCTTGCCCG 1408
Qy	841	TCGTGTCAGGAATGTCAAGTGTGATGCCAAATGCAAAAAGAGCGGGAAGAAAGAAAGA 900
Db	1409	TCGTGTCAGGAATGTCAAGTGTGATGCCAAATGCAAAAAGAGCGGGAAGAAAGAAAGC 1468
Qy	901	AAGTGACCAACATAATCTCATTTTGTATGAGGAAGATGAGCGAACTCTGGGGAGCTGT 960
Db	1469	AAGTGACCAACATAATCTCATTTTGTATGAGGAAGATGAGCAGAACTCTGGGGAGCTGT 1528
Qy	961	TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA 1020
Db	1529	TTAAAAAGACACCTGGGGCAGGGGAGAGCTCAGAGGACAACTCCGACCGCTCTCTGTCA 1588
Qy	1021	ATATCATGTCCGCTTTGAAGAGCCCTTCGGGCTTAATCTCCAAATGGAA-TCAGAGCAGCA 1079
Db	1589	ATATCATGTCCGCTTTGAAGAGCCCTTCGGGCGAACTCTCAATGGAAGTCAAGCAGCA 1648



PUBMED	12477932	
REFERENCE	2 (bases 1 to 1229)	
AUTHORS	Strausberg,R.	
TITLE	Direct Submission	
JOURNAL	Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a> Tissue Procurement: Louis Staudt cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc_mgc@nhgri.nih.gov">nisc_mgc@nhgri.nih.gov</a> Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripoop,S., Thomas,P.J., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.	
FEATURES	Location/Qualifiers	
source	1..1229	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="MGC:52410 IMAGE:4765274"	
	/tissue_type="Primary B-Cells from Tonsils"	
	/clone_lib="NIH-MGC 48"	
	/lab_host="DH10B-R"	
	/note="vector: pOTB7"	
gene	1..1229	
	/gene="FLJ12363"	
	/db_xref="LocusID:84127"	
	182..931	
	/gene="FLJ12363"	
	/codon_start=1	
	/product="FLJ12363 protein"	
	/protein_id="AAH41583.1"	
	/db_xref="GI:27371338"	
	/db_xref="LocusID:84127"	
	/translation="MDERSMLPTMAAGPNSILFAINIDKDLNGSKFAPTVSDLL KESTQNTLLKSTQGVSVFRETASSAISILIKPEQTDPLPVSRNVSADAKCK ERKKQVTNLIISFDEEDQNSGDMFKKTPGAGESSEDHSSVNMISAFSPFG NSNGSQSNWKSLSLNREFYQKLDVKSIDDEDVDENEDDDVYNSGRKRRHSE SPEKNGAHSVTQAGVQWHDLSLQPLPFGFR"	
CDS	182..931	
Query Match	33.1%; Score 630.8; DB 9; Length 1229;	
Best Local Similarity	96.8%; Pred. No. 3.6e-145;	
Matches	676; Conservative 0; Mismatches 17; Indels 5; Gaps 3;	
Qy	544 AGCACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAAAGGTCCAGTATGCTTCCT	603
Db	152 AGTACTTTTATGAAGACTGGTCTTTTGTGATGATGAAGAGAGGTCCAGTATGCTTCCT	211
Qy	604 ACCATGGCAGAGTCTGAATCCATCTCTTTGGATTAACTGACAAAGGATTG	663
Db	212 ACCATGGCAGAGTCCGAATCCATCTCTTTGGATTAACTGACAAAGGATTG	271
Qy	664 AACGGGAGAGTAAAGTTGGTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGAC	723
Db	272 AACGGGAGAGTAAAGTTGGTCCACCGTTTTCAGACCTCTTAAAGGAGTCAACGAGAA	331

Qy	724 GTGACCTCCTTGTCTGAAGGAGTCCAGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA	783
Db	332 GTCA---CCTTGTCTGAAGGAGTCCAGCAAGGAGTGAGCAGCCTGTTTCAGGGAGATCACA	388
Qy	784 GCCTCCTTCGCGCTCTCCATCTCATCAAACTGAACAGAGAGACCGA-CCCTTGGCTGTC	842
Db	389 GCCTCCTTCGCGCTCTCCATCTCATCAAACTGAACAGAGAGACCGA-CCCTTGGCTGTC	448
Qy	843 GTGTCCAGGAATGTCAAGTGTCTGATGCCAATGCAAAAGAGCGGAGAAAGAAAAGAAA	902
Db	449 GTGTCCAGGAATGTCAAGTGTCTGATGCCAATGCAAAAGAGCGGAGAAAGAAAAGCAA	508
Qy	903 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACGTGTTT	962
Db	509 GTGACCAACATAATCTCATTTGATGATGAGGAAGATGAGCAGAACTCTGGGGACATGTTT	568
Qy	963 AAAAAACACCTGGGGCAGGGAGAGGTGAGAGGACAACTCCGACCGCTCCTCTGTCAAT	1022
Db	569 AAAAAACACCTGGGGCAGGGAGAGGTGAGAGGACAACTCCGACCGCTCCTCTGTCAAT	628
Qy	1023 ATCATGTCGCGCTTTGAAAGCCCTTCGGGCTTAACCTCCCAATGAA-TCAGAGCAGCAAC	1081
Db	629 ATCATGTCGCGCTTTGAAAGCCCTTCGGGCTTAACCTCCCAATGAA-TCAGAGCAGCAAC	688
Qy	1082 TCATGGAAAAATTGATTCCTCTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTTGATGTG	1141
Db	689 TCATGGAAAAATTGATTCCTCTCTTTGAAACGGGAGTTTGGGTACCAAGAGCTTGATGTG	748
Qy	1142 AAAAGCATCGATGATGAAGATGAGGATGGAATGAAACGAGAGTACCTGTATGGAACATCATCA	1201
Db	749 AAAAGCATCGATGATGAAGATGAGGATGGAATGAAACGAGAGTACCTGTATGGAACATCATCA	808
Qy	1202 GGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA	1239
Db	809 GGAAGGAAGCACAGGGGCCACTCGGAGTCGCCCGAGAA	846

Search completed: June 10, 2005, 14:22:28  
Job time : 5427.24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:57:41 ; Search time 47.1575 Seconds  
(without alignments)  
2390.553 Million cell updates/sec

Title: US-09-155-676B-7  
Perfect score: 5052  
Sequence: 1 MAVMEMACPGAGSAGVQK.....PDGSFAWSRWKVGQLENRP 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 557933 seqs, 119041758 residues

Total number of hits satisfying chosen parameters: 557933

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:\*

- 1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US11\_NEW\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5052	100.0	947	5	US-09-155-676B-7
2	5052	100.0	947	5	US-09-155-676B-14
3	5040	99.8	947	8	US-60-659-397-1471
4	5040	99.8	947	8	US-60-659-397-1472
5	391.5	7.7	1673	8	US-60-643-717-14226
6	367	7.3	256	7	US-11-113-424-183
7	360	7.1	467	8	US-60-659-397-1644
8	360	7.1	475	6	US-10-940-774A-10456
9	359	7.1	467	5	US-09-155-676B-13
10	358	7.1	351	1	PCT-US04-30360-157
11	358	7.1	351	6	US-10-941-635-157
12	355.5	7.0	467	5	US-09-155-676B-12
13	355.5	7.0	672	5	US-09-155-676B-19
14	355.5	7.0	672	5	US-09-608-890A-2
15	355.5	7.0	1593	5	US-09-608-890A-4
16	353	7.0	619	5	US-09-608-890A-6
17	352	7.0	619	5	US-09-608-890A-8
18	349	6.9	1722	8	US-60-643-717-11386
19	346	6.8	1495	6	US-10-450-763-34736
20	343.5	6.8	614	8	US-60-643-717-691
21	341	6.7	1320	8	US-60-643-717-10799
22	338.5	6.7	611	8	US-60-643-717-689
23	337.5	6.7	314	1	PCT-US04-30360-120
24	337.5	6.7	314	6	US-10-941-635-120
25	325	6.4	1478	5	US-09-155-676B-18

26	325	6.4	1478	8	US-60-643-717-9386	Sequence 9386, Ap
27	325	6.4	1478	8	US-60-643-717-16140	Sequence 16140, A
28	325	6.4	1478	8	US-60-643-717-16764	Sequence 16764, A
29	324.5	6.4	1612	7	US-11-097-143-41304	Sequence 41304, A
30	322	6.4	1315	1	PCT-US03-27894-4	Sequence 4, Appli
31	318.5	6.3	1955	8	US-60-643-717-11116	Sequence 11116, A
32	315.5	6.2	1423	8	US-60-643-717-8279	Sequence 8279, Ap
33	313	6.2	624	6	US-10-940-774A-10942	Sequence 10942, A
34	313	6.2	626	1	PCT-US03-10870-451	Sequence 451, App
35	313	6.2	626	5	US-09-608-890A-10	Sequence 10, Appl
36	307	6.1	1832	8	US-60-643-717-11014	Sequence 11014, A
37	305.5	6.0	525	6	US-10-491-545A-25	Sequence 25, Appl
38	304.5	6.0	522	6	US-10-703-032-136408	Sequence 136408,
39	304.5	6.0	965	7	US-11-090-997-508	Sequence 508, App
40	304.5	6.0	1490	1	PCT-US03-10870-912	Sequence 912, App
41	304.5	6.0	2139	7	US-11-097-143-13971	Sequence 13971, A
42	303	6.0	2523	6	US-10-130-978-60	Sequence 60, Appl
43	301.5	6.0	968	1	PCT-US04-42360-1857	Sequence 1857, Ap
44	301.5	6.0	968	6	US-10-940-774A-6680	Sequence 6680, Ap
45	301.5	6.0	974	7	US-11-090-997-1594	Sequence 1594, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-155-676B-7  
; Sequence 7, Application US/09155676B  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR  
; FILE REFERENCE: WALLACH=21  
; CURRENT APPLICATION NUMBER: US/09/155.676B  
; CURRENT FILING DATE: 1999-01-04  
; PRIOR APPLICATION NUMBER: PCT/IL97/00117  
; PRIOR FILING DATE: 1997-04-01  
; PRIOR APPLICATION NUMBER: IL 117800  
; PRIOR FILING DATE: 1996-04-02  
; PRIOR APPLICATION NUMBER: IL 119133  
; PRIOR FILING DATE: 1996-08-26  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 7  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-155-676B-7

Query Match 100.0%; Score 5052; DB 5; Length 947;  
Best Local Similarity 100.0%; Pred. No. 6.9e-231;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db	1	MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Qy	61	VITKGTAKGSEAGPAAISIIAQACENQSFPTFSERIFIAGSKOYSQSESLDQIPNN 120
Db	61	VITKGTAKGSEAGPAAISIIAQACENQSFPTFSERIFIAGSKOYSQSESLDQIPNN 120
Qy	121	VAHATEGMARVCKGKRKSRKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Db	121	VAHATEGMARVCKGKRKSRKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
Qy	181	DESLPGAPYVNTPQFTKPLKEGGLGQCFKQLEGRLPALPRSELHKLISPLQCLNHW 240
Db	181	DESLPGAPYVNTPQFTKPLKEGGLGQCFKQLEGRLPALPRSELHKLISPLQCLNHW 240



```
QY 241 KLHPDQGGPLPHTPPYSLRPLPDPHPPLPQPKPHPLESLGKLACVDSOKPLPDPH 300
Db 241 KLHPDQGGPLPHTPPYSLRPLPDPHPPLPQPKPHPLESLGKLACVDSOKPLPDPH 300
QY 301 LSKLACVDSOKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSOKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRREPSKPTEDNEGVLITTEKLKPDYVEYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AARGSRREPSKPTEDNEGVLITTEKLKPDYVEYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIRVPLYGAVREGPWNNIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIRVPLYGAVREGPWNNIFMELLEGGSLGOL 480
QY 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASSEPPVREIIPSCAPLTAQAIQEGLRKEPIHRSVAELGGKVNRAALQOQVG 660
Db 601 RGPLCLKIASSEPPVREIIPSCAPLTAQAIQEGLRKEPIHRSVAELGGKVNRAALQOQVG 660
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
QY 721 PEPNKSPPILTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELEFLN 780
Db 721 PEPNKSPPILTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELEFLN 780
QY 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSSQAEARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
QY 901 LVTXKQGPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947
Db 901 LVTXKQGPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947

RESULT 2
US-09-155-676B-14
; GENERAL INFORMATION: US/09155676B
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAP), THEIR
; PREPARATION AND USE
; FILE REFERENCE: WALLACH-21
; CURRENT APPLICATION NUMBER: US/09/155,676B
; CURRENT FILING DATE: 1999-01-04
; PRIOR APPLICATION NUMBER: PCT/IL97/00117
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: IL 117800
; PRIOR FILING DATE: 1996-04-02
; PRIOR APPLICATION NUMBER: IL 119133
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 947
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-676B-14

Query Match      100.0%; Score 5052; DB 5; Length 947;
Best Local Similarity 100.0%; Pred. No. 6.9e-231;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACGAPGSAVGQOKELPKPKETPTPLPKGQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACGAPGSAVGQOKELPKPKETPTPLPKGQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKEGSAGAAISITAAACENSQESPTFSERIFIAGSKQYSQSSLSQIIPNN 120
Db 61 VITKGTAKEGSAGAAISITAAACENSQESPTFSERIFIAGSKQYSQSSLSQIIPNN 120
QY 121 VAHATEGKMARVCWKGRSKARKKKSSKSLAHAGVALAKPLPRPEQSSCTIPVQE 180
Db 121 VAHATEGKMARVCWKGRSKARKKKSSKSLAHAGVALAKPLPRPEQSSCTIPVQE 180
QY 181 DESPLGAPYVRNTPOFTKPLKEPGLGOLCFKOLGELRPAUPRSEHLKLIISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLGOLCFKOLGELRPAUPRSEHLKLIISPLQCLNHVW 240
QY 241 KLHHPDQGGPLPHTPPYSLRPLPDPHPPLPQPKPHPLESLGKLACVDSOKPLPDPH 300
Db 241 KLHHPDQGGPLPHTPPYSLRPLPDPHPPLPQPKPHPLESLGKLACVDSOKPLPDPH 300
QY 301 LSKLACVDSOKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSOKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRREPSKPTEDNEGVLITTEKLKPDYVEYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AARGSRREPSKPTEDNEGVLITTEKLKPDYVEYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIRVPLYGAVREGPWNNIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIRVPLYGAVREGPWNNIFMELLEGGSLGOL 480
QY 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYYLGQALEGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASBPPVREIIPSCAPLTAQAIQEGLRKEPIHRSVAELGGKVNRAALQOQVG 660
Db 601 RGPLCLKIASBPPVREIIPSCAPLTAQAIQEGLRKEPIHRSVAELGGKVNRAALQOQVG 660
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPRAEETTGAPKLQPLPPE 720
QY 721 PEPNKSPPILTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELEFLN 780
Db 721 PEPNKSPPILTLSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQLOLELEFLN 780
QY 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSSQAEARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900
QY 901 LVTXKQGPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947
Db 901 LVTXKQGPVRYDMEVPDSDGIDLQCTLAPDGSFAWSWRVKHGOLENRP 947
```



```
RESULT 3
US-60-659-397-1471
; Sequence 1471, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1471
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1471

Query Match      99.8%; Score 5040; DB 8; Length 947;
Best Local Similarity 99.8%; Pred. No. 2.6e-230;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIPIAGSKQYSQESLDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIPIAGSKQYSQESLDQIPNN 120
QY 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
DB 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
QY 181 DESPLGAPYVRNTPTQFTKPLKEPGLGQCFKQGLGELRPAIPRSELHKLISPLQCLNHV 240
DB 181 DESPLGAPYVRNTPTQFTKPLKEPGLGQCFKQGLGELRPAIPRSELHKLISPLQCLNHV 240
QY 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPPLQPKWHPLESFLGKLACVDSQKPLDPH 300
DB 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPPLQPKWHPLESFLGKLACVDSQKPLDPH 300
QY 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKFSVEEYLVAHQGSVSSQAHSLSLAKTW 360
DB 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKFSVEEYLVAHQGSVSSQAHSLSLAKTW 360
QY 361 AAGRSRREPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
DB 361 AAGRSRREPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPMELLEGGSLGQL 480
DB 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPMELLEGGSLGQL 480
QY 481 VKEQGLCEDRALYVLGALGLEYLHSHRRILHGDVKADNVLSSDGSAAALCDFGHAVC 540
DB 481 VKEQGLCEDRALYVLGALGLEYLHSHRRILHGDVKADNVLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGSLTLTGDIYIPGTEHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
DB 541 LQPDGLGSLTLTGDIYIPGTEHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASPPPPVREIPSPCAPITAQAIQGLRKEPIHRVSAEALGGKVNRAALQVGG 660
DB 601 RGPLCLKIASPPPPVREIPSPCAPITAQAIQGLRKEPIHRVSAEALGGKVNRAALQVGG 660
QY 661 LKSPWRGEYKEPRHPHPNQAQYHQLHAQPRELSPRAPCPRAEETTGAPKLPPLPPE 720
DB 661 LKSPWRGEYKEPRHPHPNQAQYHQLHAQPRELSPRAPCPRAEETTGAPKLPPLPPE 720

RESULT 4
US-60-659-397-1472
; Sequence 1472, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1472
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1472

Query Match      99.8%; Score 5040; DB 8; Length 947;
Best Local Similarity 99.8%; Pred. No. 2.6e-230;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
DB 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIPIAGSKQYSQESLDQIPNN 120
DB 61 VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSERIPIAGSKQYSQESLDQIPNN 120
QY 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
DB 121 VAHATEGMARVCWKGRSKARKKRSKSLAHAGVALAKPLPRTEOESCTIPVOE 180
QY 181 DESPLGAPYVRNTPTQFTKPLKEPGLGQCFKQGLGELRPAIPRSELHKLISPLQCLNHV 240
DB 181 DESPLGAPYVRNTPTQFTKPLKEPGLGQCFKQGLGELRPAIPRSELHKLISPLQCLNHV 240
QY 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPPLQPKWHPLESFLGKLACVDSQKPLDPH 300
DB 241 KLHPDQGGPLPLTPHPFPYSRLPHPPFHPPLQPKWHPLESFLGKLACVDSQKPLDPH 300
QY 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKFSVEEYLVAHQGSVSSQAHSLSLAKTW 360
DB 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKFSVEEYLVAHQGSVSSQAHSLSLAKTW 360
QY 361 AAGRSRREPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
DB 361 AAGRSRREPSPKTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPMELLEGGSLGQL 480
DB 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVLVYAVREGPWNIIPMELLEGGSLGQL 480
QY 481 VKEQGLCEDRALYVLGALGLEYLHSHRRILHGDVKADNVLSSDGSAAALCDFGHAVC 540
DB 481 VKEQGLCEDRALYVLGALGLEYLHSHRRILHGDVKADNVLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGSLTLTGDIYIPGTEHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
DB 541 LQPDGLGSLTLTGDIYIPGTEHMAPEVVLGRSCDAKVDWVSSCCMMLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASPPPPVREIPSPCAPITAQAIQGLRKEPIHRVSAEALGGKVNRAALQVGG 660
DB 601 RGPLCLKIASPPPPVREIPSPCAPITAQAIQGLRKEPIHRVSAEALGGKVNRAALQVGG 660
QY 661 LKSPWRGEYKEPRHPHPNQAQYHQLHAQPRELSPRAPCPRAEETTGAPKLPPLPPE 720
DB 661 LKSPWRGEYKEPRHPHPNQAQYHQLHAQPRELSPRAPCPRAEETTGAPKLPPLPPE 720
```



```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
US-11-113-424-183

Query Match          7.3%; Score 367; DB 7; Length 256;
Best Local Similarity 37.9%; Pred. No. 2.2e-10;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

Qy      406 LGRSGFGEVHRMEDKOTGFCAVKKKVRLEVFRAE-----ELMACAGLTSPRIPLYGA 458
       |||::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      7 LGKAFGKYVLARDKTKGLVAIKVIKEKLKKRERIIRIILKUKLDHNIVKLVDY 66

Qy      459 VREGPWNIIFMELLEGGSLGOLVKEOCLPEDRALYYLGOALEGLEYLHSRRILHGDVKA 518
       :   :   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      67 FEDDDKLVNMEYCEGGDLFDLLKKRGRLSEARFVARQILSALBYLSOGIIHRDLKP 126

Qy      519 DNVLSSDGSAAALCDRHAVCILOPDLGSKSLTGYITGTETHMAPEVILGRSCDAKD 578
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      127 ENILLSDG-HVKLAFLAKQLDSGG---TLTT--TFVGTPHYMAPEVLLKGYGKAVD 180

Qy      579 WSSCCMWMLHLNGCHPWTOFFRG-----PLCLKIASPPPVREIPP---SCAPLTAQAI 630
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      181 IWSLGVLIELTGKPP----PFDDOLLALFKIGKPPP---FPDPWKISPAEKDLI 233

Qy      631 QEGLRKEPIHRVSAAE 646
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      234 KKLKVDPCKRLTAEE 249


RESULT 7
US-60-659-397-1644
; Sequence 1644, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1644
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-659-397-1644

Query Match          7.1%; Score 360; DB 8; Length 467;
Best Local Similarity 27.2%; Pred. No. 9e-10;
Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;

Qy      332 SVEEYLVALQGVSQAHLSTLAKTWAAGRSRREPSPKTENEGVLLTEKLKPDY 391
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      80 TVEDLL-AFANHISNTAKHFY-----GORPOE-----SGILLNNVITPQNG 119

Qy      392 EYREE-----VHWATHQLRLG-----RGSFGEVHRMEDKQTGFQCALKKRVLEVFAEEL 441
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      120 RYQIDSVDLLIPWKLTYRNIGSDFIPIRGAFGVLYLAQDIKTKRMACKLIPIVDQPKPSDV 179

Qy      442 MACAGLTSPRIPLYGAVREGPWNIIFMELLEGGSLGOLVKEOGLCPEDRALYYLGOALE 501
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      180 EIQACFRHENIAELYGAVLWGNETVHLFMAGEGGSVLEKLES CGPMREFEIITVTXHVLUK 239

Qy      502 GLEYLHSRRILHGDKADNVLLSSDGSHAAALCDFGHAVCLQDPDLGKSLLTDGYTP----- 557
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      240 GLDFLHKSVIHHDIKPSNIVFMS--TKAVLVDFGLSVQMTE-----VYFPKDLR 288

Qy      557 GTETHMAPEVILGRSCDAKDQVWSSCCMWMLHLNGCHPWTOFF---RGPLCLKIA-SEPP 613
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      297 GTIEYSPEVILCRGSTKADIYSLGATIHHMTGTPPWVKRYPRSAVPYSYLIITHKQAP 356

Qy      614 PVREIPPCAPLTAQAIQEGLRKPEIHRVSAAEELGKNRNLQQVGGKLPWRGEYKEPR 673
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      357 PLEDIADDCCPGMRRELIEASLERPNHPRAADL-----LK----- 392

Qy      674 HPPNQANYHOTLHAOPRELSPRAPGPRAEETTGRAPKLQDPPLPEPPEPNKSPPLTUS 733
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      393 -----HEALN-PREDQPCR-----QSILDSALLERKELLRSKELELPENTADSSCTGS 439



```

QY 734 KEESGM 739  
Db 440 TERSEM 445

RESULT 9

US-09-155-676B-13

; Sequence 13, Application US/09155676B

; GENERAL INFORMATION:

; APPLICANT: WALLACH, David

; APPLICANT: MALININ, Nikolai

; APPLICANT: BOLDIN, Mark

; APPLICANT: KOVALENKO, Andrei

; APPLICANT: METT, Igor

; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR

; PREPARATION AND USE

; FILE REFERENCE: WALLACH-21

; CURRENT APPLICATION NUMBER: US/09/155,676B

; CURRENT FILING DATE: 1999-01-04

; PRIOR APPLICATION NUMBER: PCT/IL97/00117

; PRIOR FILING DATE: 1997-04-01

; PRIOR APPLICATION NUMBER: IL 117800

; PRIOR FILING DATE: 1996-04-02

; PRIOR APPLICATION NUMBER: IL 119133

; PRIOR FILING DATE: 1996-08-26

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 13

; LENGTH: 467

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-155-676B-13

Query Match 7.1%; Score 359; DB 5; Length 467;

Best Local Similarity 27.2%; Pred. No. 1e-09; Indels 88; Gaps 14;

Matches 116; Conservative 69; Mismatches 154; Indels 88; Gaps 14;

QY 332 SVEEYLHALQSVSSQAHSLTSLAKTWAAGRSRSPKTEDNEGVLLEKLPVDY 391

Db 80 TVEDLL--AFANHSNTAKHFY-----GQRPQE-----SGILLNNVITPQNG 119

QY 392 EYREE-----VHWATHQLRLG-----RGSFGVHRMEDKQTGFQCAVKVRLEVFRASEL 441

Db 120 RYQIDSDVLLIPWKLTYNIGSDFTPRGAFGVYLAQDIKTKRMACKLIPVDQFKPSDV 179

QY 442 MACAGLTSPRIVELYGAVREGPWNIEMELLEGGSLGOLVKEQCLPEDRALYVLGQALE 501

Db 180 ETQACFRHENIAELYGAVLWGTVHLFMEAGEGSVLEKLESCGPMRFEIITWTKHVLK 239

QY 502 GLEYLHRRILHGDVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLTGDYIP--- 557

Db 240 GLDFLHKKVIHHDIKPSNIVFWS--TKAVLVDFGLSV-----QMTEDVYFVKDL 287

QY 558 -GTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFF---RGPLCLKTA-SEP 612

Db 288 RTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKRYPRSAYPEIYIHKQA 347

QY 613 PVRBIPSPCAPLTAQAIQEGRLKEPIHRSVAEELGGKVNRAALQOVGGLKSPWRGEYKEP 672

Db 348 PPLEDIADDCSPGMRELIEASLERPNHRPRAADL-----LK----- 384

QY 673 RHPPNQANYHTLHAQRELSAPRPAETTRAPKLOPPLPEPPPEPNKSPPLTL 732

Db 385 -----HEALN-PPREDQPRC---QSLDSALLERKRLLSRKELELPENIADSSCTG 430

QY 733 SKEESGM 739

Db 431 TERSEM 437

RESULT 10

PCT-US04-30360-157

; Sequence 157, Application PC/TUS0430360

; GENERAL INFORMATION:

; APPLICANT: PLEXIKON, INC.

; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT

; FILE REFERENCE: 039363-1703

; CURRENT APPLICATION NUMBER: PCT/US04/30360

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/503,277

; PRIOR FILING DATE: 2003-09-15

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 157

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Unknown Organism

; FEATURE:

; OTHER INFORMATION: Description of Unknown Organism: Mammalian

; OTHER INFORMATION: protein sequence

PCT-US04-30360-157

Query Match 7.1%; Score 358; DB 1; Length 351;

Best Local Similarity 27.7%; Pred. No. 8.2e-10;

Matches 113; Conservative 68; Mismatches 151; Indels 76; Gaps 13;

QY 350 AHSLSLAKTWAAGRSRSPKTEDNEGVLLEKLPVDYVEYREE-----VHWATHQL 404

Db 1 ANHSNTAKHP--YQORPQE-----SGILLNNVITPQNGRYQIDSDVLLIPWKLTYSR 50

QY 405 RLG-----RGSFGVHRMEDKQTGFQCAVKVRLEVFRAEELMACAGLTSPRIVELYGAV 459

Db 51 NIGSDFIPIRGAFGVYLAQDIKTKRMACKLIPVDQFKPSDVEIQACFRHENIAELYGAV 110

QY 460 REGPWNIEMELLEGGSLGOLVKEQCLPEDRALYVLGQALEGLEYLHRRILHGDVKAD 519

Db 111 LWGETVHLFMEAGEGSVLEKLESCGPMRFEIITWTKHVLKGLDFLHKKVIHHDIKPS 170

QY 520 NVLLSSDGSAAALCDFGHAVCLQPDGLGKSLTGDYIP---GTETHMAPEVVLGRSCDA 575

Db 171 NIVFWS--TKAVLVDFGLSVQMTED-----VYFPKDLRGTEIYMSPEVILCRGHST 219

QY 576 KVDVWSSCCMLHMLNGCHPWTOFF---RGPLCLKIA-SEPPVPRIPSPCAPLTAQAIQ 631

Db 220 KADIYSLGATLIHMOTGTPPWVKRYPRSAYPEIYIHKQAPPLIEDIADDCSPGMRELIE 279

QY 632 EGLRKEPIHRSVAEELGGKVNRAALQOVGGLKSPWRGEYKEPDRHPPNQANYHTLHAQPR 691

Db 280 ASLERPNHRPRAADL-----LK-----HEALN-PPR 305

QY 692 ELSRAPGRPAETTRAPKLOPPLPEPPPEPNKSPPLTLKSKEESGM 739

Db 306 EDQPRCTS---LDSALLERKRLLSRKELELPENIADSSCTGSTEESEM 350

RESULT 11

US-10-941-635-157

; Sequence 157, Application US/10941635

; GENERAL INFORMATION:

; APPLICANT: ARTIS, DEAN R.

; APPLICANT: BREMER, RYAN E.

; APPLICANT: GILLETTE, SAMUEL J.

; APPLICANT: HURT, CLARENCE R.

; APPLICANT: IBRAHIM, PRABHA L.

; APPLICANT: ZUCKERMAN, REBECCA L.

; TITLE OF INVENTION: MOLECULAR SCAFFOLDS FOR KINASE LIGAND DEVELOPMENT

; FILE REFERENCE: 039363-1702

; CURRENT APPLICATION NUMBER: US/10/941,635

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/503,277

; PRIOR FILING DATE: 2003-09-15

; NUMBER OF SEQ ID NOS: 167

; SOFTWARE: PatentIn Ver. 3.2

; SEQ ID NO 157

; LENGTH: 351

; TYPE: PRT

```
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Mammalian
; OTHER INFORMATION: protein sequence
US-10-941-635-157

Query Match          7.1%; Score 358; DB 6; Length 351;
Best Local Similarity 27.7%; Pred. No. 8.2e-10;
Matches 113; Conservative 68; Mismatches 151; Indels 76; Gaps 13;

QY 350 AHSLSLAKTAAAGSRSPKTEKNEGVLITKLPVDVEYRES-----VHWATHQL 404
DB 1 ANHISNTAKHP--YQORPQE-----SGILLNNVITPQGRYQIDSDVLLIPWKLYR 50

QY 405 RLG-----RGSFGEVHRMEDKQTQCQCAVKVRLVFRABELMACAGLTSPRIVLYGAV 459
DB 51 NIGSDFIPRGAGFKVYLAQDIKTKRMACKLIPVDQPKPSDVEIQACFRHENIAELYGAV 110

QY 460 REGPWNINFMELLEGSGQLVKEQGLPDEBALYYLQALLEGLEYLHSRILHGDVKAD 519
DB 111 LWGETVHLUFMEAGEGGSVLEKLESCGPMREFEIIWVTRKVLKGLDFLHSHKVKVHHDIKPS 170

QY 520 NVLLSSDGSAAALCDFGHAVCLQPDGLGKSLITGDYIP-----GTETHMAPEVVLGRSCDA 575
DB 171 NIVFMS---TKAVLVDFGLSVQMTED-----VFPKDLRGTEIYMSPEVILCRGHST 219

QY 576 KVDVSSCCMLHMLNGCHPWTQFF---RGPLCLKIA-SEPPVPVREIPSPSCAPLTAQAIQ 631
DB 220 KADIYSLGATLIHMQTGTPPWVKYPRSAVPSYLYIHKQAPPLEDIADDCSPGMRELIE 279

QY 632 EGLRKEPIHRVSAELGKGVNRALQOQVGLKSPWRGEYKEPRHPPPNQANYHOTLHAQPR 691
DB 280 ASLERNPNRPRADL-----LK-----HEALN-PPR 305

QY 692 ELSRAPCPRAEETTGAPKLQPLPEPEPNKSPPLTLKSRESGM 739
DB 306 EQDPRCTS---LDSALLERKLLSRKELELPENIADSSCTGSTBESEM 350

RESULT 12
US-09-155-676B-12
; Sequence 12, Application US/09155676B
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
; FILE REFERENCE: WALLACH=21
; CURRENT APPLICATION NUMBER: US/09/155.676B
; PRIOR FILING DATE: 1999-01-04
; PRIOR FILING DATE: 1997-04-01
; PRIOR FILING DATE: 1996-04-02
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 12
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-676B-12

Query Match          7.1%; Score 356.5; DB 5; Length 467;
Best Local Similarity 29.2%; Pred. No. 1.3e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY 379 GVLITKLPVDVEYREE-----VHWATHQLRLG-----RGSFGEVHRMEDKQTGFQCAV 428
DB 1 ANHISNTAKHP--YQORPQE-----SGILLNNVITPQGRYQIDSDVLLIPWKLYR 50

QY 350 AHSLSLAKTAAAGSRSPKTEKNEGVLITKLPVDVEYRES-----VHWATHQL 404
DB 1 ANHISNTAKHP--YQORPQE-----SGILLNNVITPQGRYQIDSDVLLIPWKLYR 50

QY 405 RLG-----RGSFGEVHRMEDKQTQCQCAVKVRLVFRABELMACAGLTSPRIVLYGAV 459
DB 51 NIGSDFIPRGAGFKVYLAQDIKTKRMACKLIPVDQPKPSDVEIQACFRHENIAELYGAV 110

QY 460 REGPWNINFMELLEGSGQLVKEQGLPDEBALYYLQALLEGLEYLHSRILHGDVKAD 519
DB 111 LWGETVHLUFMEAGEGGSVLEKLESCGPMREFEIIWVTRKVLKGLDFLHSHKVKVHHDIKPS 170

QY 520 NVLLSSDGSAAALCDFGHAVCLQPDGLGKSLITGDYIP-----GTETHMAPEVVLGRSCDA 575
DB 171 NIVFMS---TKAVLVDFGLSVQMTED-----VFPKDLRGTEIYMSPEVILCRGHST 219

QY 576 KVDVSSCCMLHMLNGCHPWTQFF---RGPLCLKIA-SEPPVPVREIPSPSCAPLTAQAIQ 631
DB 220 KADIYSLGATLIHMQTGTPPWVKYPRSAVPSYLYIHKQAPPLEDIADDCSPGMRELIE 279

QY 632 EGLRKEPIHRVSAELGKGVNRALQOQVGLKSPWRGEYKEPRHPPPNQANYHOTLHAQPR 691
DB 280 ASLERNPNRPRADL-----LK-----HEALN-PPR 305

QY 692 ELSRAPCPRAEETTGAPKLQPLPEPEPNKSPPLTLKSRESGM 739
DB 306 EQDPRCTS---LDSALLERKLLSRKELELPENIADSSCTGSTBESEM 350

RESULT 13
US-09-155-676B-19
; Sequence 19, Application US/09155676B
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR
; FILE REFERENCE: WALLACH=21
; CURRENT APPLICATION NUMBER: US/09/155.676B
; PRIOR FILING DATE: 1999-01-04
; PRIOR FILING DATE: 1997-04-01
; PRIOR FILING DATE: 1996-04-02
; PRIOR FILING DATE: 1996-08-26
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 19
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-155-676B-19

Query Match          7.0%; Score 355.5; DB 5; Length 672;
Best Local Similarity 23.6%; Pred. No. 2.2e-09;
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;

QY 77 AISIIAQAECENSQEFSPFTSERIFIAGSKQYQSESLDQIPNNVAHATEGMARVCWK 136
DB 58 SLQAVATSCLENSLSLEHTVHREKTGSLNTRLSASSEDISDLAGVSVGLPSSTTTEQ 117

QY 137 KR---RSKARKRKXKSKSLAHAGVALAKPLPTPEQESCTIPVQEDSPGAPYVRNT 193
DB 118 PKPAVQTKGRPHSQCLNSPLSHA--QLMPFAPSAPCSSAPSVP---DIS-----KHR 165

QY 194 PQFTKPLKEPCLGQLCFKQJGEGRLRPAIPRSELHKLISPLQ--CLNHVWKLHHPDGGPL 251
DB 166 PQAFVPCPKPS-----ASPQTQ-RKFSLQFORNCSEH-----RDSQOL 202

QY 252 PLPTHPPFYRLPHFPFPHLPQPMKPHPLESFLGKLACVDSQKPLPDPHLSKLA----- 305
DB 203 ---SPVQTGRPPSSNIHRPKPSRPV-----GTSKLGATKSSMTILDLSA 248

QY 306 --CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEVLVHALQGSVSS-- 347
DB 1 ANHISNTAKHP--YQORPQE-----SGILLNNVITPQGRYQIDSDVLLIPWKLYR 50
```

Db 249 SRCDSFGGNGSGNAVIPSDETFTVPEVKCRDLVNTLNSSIEDLLEASMPSSDITVT 308  
QY 348 -----SQAHSLTSLAKTWAARSR---EPSP 372  
Db 309 FKESEVAVLSPEKAENDTYKDDVNNHKKCKEKEAEEAALAMAMASQDALPIVQOL 368  
QY 373 KTEDNEGVLLEKLP-----VDYEREVEVHATHQRLGRGSGFEVHRMEDKQTGF 424  
Db 369 QVENGEDIIIIQODTETPLPGHTKAKOPYREDAEWLKGQ-QTGLGAFSSCYQAQDVGTGT 427  
QY 425 QCAVKV-----RLEVFR--BELMACAGLTSPRIVLYGAVREGPWNIEMELLE 473  
Db 428 LMAVKQVTVYRNTSSQEVEVEALREERMMGHLPNIIRMLGATCEKSNYNLFIEWMA 487  
QY 474 GSGLGOLVKEQCLPEDRALYLGQALEGLYLSRILHGDVKNVLLSSDGSAAALC 533  
Db 488 GGSVAHLLSKYGAFKESVVINYTEQLRGLSYLHENQIIHRDVKGANLLIDSTGQRLRIA 547  
QY 534 DFGHAVCLQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593  
Db 548 DFGAAARLASKGTGAGEFQGLL-GTIAFMAPEVLGQQYGRSCDVWSVGCALIEMACAK 606  
QY 594 HPWTQFRG---PLCLKIASE-----PPVREIIPSCAPLTAQ--AIQELGRKEP 638  
Db 607 PPWNAEKHNHLALIFKIASATTAPSIPLSHLSPGLRDVAVRCLELQPDQRPSPRELLKHP 666  
QY 639 IHRVS 643  
Db 667 VERTT 671

RESULT 14  
US-09-608-890A-2  
; Sequence 2, Application US/09608890A  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To  
; TITLE OF INVENTION: External Signals  
; FILE REFERENCE: CPI-004DVC3CN  
; CURRENT APPLICATION NUMBER: US/09/608,890A  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 08/440,421  
; PRIOR FILING DATE: 1993-05-15  
; PRIOR APPLICATION NUMBER: 08/323,460  
; PRIOR FILING DATE: 1994-10-14  
; PRIOR APPLICATION NUMBER: 08/410,602  
; PRIOR FILING DATE: 1995-03-24  
; PRIOR APPLICATION NUMBER: 08/472,934  
; PRIOR FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 672  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-608-890A-2

Query Match 7.0%; Score 355.5; DB 5; Length 672;  
Best Local Similarity 23.6%; Pred. No. 2.2e-09;  
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;  
QY 77 AISIIAQAECENSQBFSTFSERIFAGSKQYSQSESLDQIPNNVAHATEGMARVCWK 136  
Db 58 SLQAVAPTSCLENSLSLEHTVHREKTKGLSATRLSASSEDISDRLAGVSVGLPSTTTEQ 117  
QY 137 KR---RSKARKKRRKSKSLAHAGVALAKPLPRTPEQSCCTIPVQEDSEPLGAPVNT 193  
Db 118 PKPAVQTKGRPHSQCLNSPLSHA--OLMFPAPSAPCSSAPSPV---DIS-----K 165  
QY 194 PQTCKPLKEPGLQCLCFKQGLBGLRPLPRSELHLKLSPLQ--CLNHWKLLHPDGGPL 251

Db 166 PQAFVPCIPS-----ASPQTQ-RKFSLQFQNCSEH-----RDSOL 202  
QY 252 PLPTHTPPYSLRPLPHPPHLPQWPKHPLESFLGKLACVDSQKPLPDPHLSKLA----- 305  
Db 203 ---SVFTQSRPPPPSNIRPKPSRPV-----GSTSKLGDATKSMTLDLGS 248  
QY 306 --CVDS-----PKPLPGPH-----LEPSC--LSRGAHEKFSVEYLVAHQGVSS-- 347  
Db 249 SRCDSFGGNGSGNAVIPSDETFTVPEVKCRDLVNTLNSSIEDLLEASMPSSDITVT 308  
QY 348 -----SQAHSLTSLAKTWAARSR---EPSP 372  
Db 309 FKESEVAVLSPEKAENDTYKDDVNNHKKCKEKEAEEAALAMAMASQDALPIVQOL 368  
QY 373 KTEDNEGVLLEKLP-----VDYEREVEVHATHQRLGRGSGFEVHRMEDKQTGF 424  
Db 369 QVENGEDIIIIQODTETPLPGHTKAKOPYREDAEWLKGQ-QTGLGAFSSCYQAQDVGTGT 427  
QY 425 QCAVKV-----RLEVFR--EELMACAGLTSPRIVLYGAVREGPWNIEMELLE 473  
Db 428 LMAVKQVTVYRNTSSQEVEVEALREERMMGHLPNIIRMLGATCEKSNYNLFIEWMA 487  
QY 474 GSGLGOLVKEQCLPEDRALYLGQALEGLYLSRILHGDVKNVLLSSDGSAAALC 533  
Db 488 GGSVAHLLSKYGAFKESVVINYTEQLRGLSYLHENQIIHRDVKGANLLIDSTGQRLRIA 547  
QY 534 DFGHAVCLQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593  
Db 548 DFGAAARLASKGTGAGEFQGLL-GTIAFMAPEVLGQQYGRSCDVWSVGCALIEMACAK 606  
QY 594 HPWTQFRG---PLCLKIASE-----PPVREIIPSCAPLTAQ--AIQELGRKEP 638  
Db 607 PPWNAEKHNHLALIFKIASATTAPSIPLSHLSPGLRDVAVRCLELQPDQRPSPRELLKHP 666  
QY 639 IHRVS 643  
Db 667 VERTT 671

RESULT 15  
US-09-608-890A-4  
; Sequence 4, Application US/09608890A  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To  
; TITLE OF INVENTION: External Signals  
; FILE REFERENCE: CPI-004DVC3CN  
; CURRENT APPLICATION NUMBER: US/09/608,890A  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 08/440,421  
; PRIOR FILING DATE: 1995-05-15  
; PRIOR APPLICATION NUMBER: 08/323,460  
; PRIOR FILING DATE: 1994-10-14  
; PRIOR APPLICATION NUMBER: 08/049,254  
; PRIOR FILING DATE: 1993-04-15  
; PRIOR APPLICATION NUMBER: 08/410,602  
; PRIOR FILING DATE: 1995-03-24  
; PRIOR APPLICATION NUMBER: 08/472,934  
; PRIOR FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1593  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-608-890A-4

Query Match 7.0%; Score 355.5; DB 5; Length 1593;  
Best Local Similarity 23.6%; Pred. No. 5.5e-09;  
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;  
QY 77 AISIIAQAECENSQBFSTFSERIFAGSKQYSQSESLDQIPNNVAHATEGMARVCWK 136

```
Db 979 SLQAVPTSCLENSLEHTVHREKTGKGLSATRLSASSEDISDRLAGVSVGLPSPSTTEQ 1038
Qy 137 KR---RSKARKRKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQEDESPFGAPYVRNT 193
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1039 PKPAVQTKGRPHSQCLNSSPLSHA--QLMFAPSAPCSAPSPV---DIS-----KHR 1086
Qy 194 POFYKPLKEPGLGQLCFQOLGEGLRPALPRSELHKLISPLQ--CLNHVWKLHHFQDGGPL 251
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1087 POAFVPCPKIPS-----ASPQO-RKFSLQFORNCSEH-----RDSQOL 1123
Qy 252 PLPTHPPFPYSLPHPPPLQPMKPHLESFLGKLACVDSOKPLPDPHLSKLA----- 305
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1124 ---SPVFTQSRPPSSNHRPKSRPVP-----GSTKLGDKATKSMTDLGSA 1169
Qy 306 --CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEYLVHALQGSVSS-- 347
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1170 SRCDDSGGGGSGNAVIPSDETVTPVEDKCRLDVNTLNSSIEDLLEASMPSSDTTIVT 1229
Qy 348 -----SQHSLTSLAKTWAARGRSR---BPSP 372
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1230 FKSEVAVLSPEKAENDDTYKDDVNHNOCKEKEEAEAEALAIAMANSASQDALPIVPOL 1289
Qy 373 KTEDEGVLLEKLKP-----VDYEYREEVHATHQLRLGRSGFEVHRMEDKOTGF 424
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1290 QVENGEIIIIQQDTPETLPQHTKAKQPYREDAEWLKGQ-QIGLGAFSSCYQAQDVGTGT 1348
Qy 425 QCAVKKV-----RLEVFRA--EELMACAGLTSRIVPLYGAVREGPWVNIFFMELLE 473
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1349 LMVAKQVTVYVNTSSEQEVEVEALREEIRMMGHLNHPNIIIRMLGATCEKSNYNLFIEWMA 1408
Qy 474 GGSGLQGVKEGCLPEDRALYYLGOALEGLEYLHSRRILHGDVXADNVLLSSDGSAAALC 533
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1409 GGSVAHLLSKYGAFKESVWINYTEQLLRGLSYLHENQIIHRDVKGANLLIDSTGQRLRIA 1468
Qy 534 DFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1469 DFGAAARLASKGTGAGEFQOQLL-GTIAFMAPEVLRGQYGRSCDVWSVGCIIEMACAK 1527
Qy 594 HPMTQFFRG---PLCLKIASE-----PPVREIPSPCAPLTAQ--AIQGLRKEP 638
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1528 PPWNAEKHSNHLALIFKIASATTAPSPHLSFGLRDVAVFCLELOQDRPPSRELLKHP 1587
Qy 639 IHRVS 643
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1588 VFRIT 1592
```

Search completed: June 7, 2005, 12:21:21  
Job time : 50.1575 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: June 7, 2005, 11:52:21 ; Search time 422.493 Seconds  
(without alignments)  
2618.037 Million cell updates/sec

Title: US-09-155-676B-7  
Perfect score: 5052  
Sequence: 1 MAVMEMACPGAGSAGVQOK.....PDGSFAMSWRVKHGLENR 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/paa/US080\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
- 21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*
- 22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*
- 23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
- 24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*
- 25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*
- 26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*
- 27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*
- 28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*
- 29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*
- 30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*
- 31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*
- 32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*
- 33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*
- 34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*
- 35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*
- 36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep.\*
- 37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5052	100.0	947	15	US-09-155-676-7
2	5052	100.0	947	15	US-09-155-676-14
3	5052	100.0	947	15	US-09-155-676A-7
4	5052	100.0	947	15	US-09-155-676A-14
5	5052	100.0	947	22	US-09-791-537-1606
6	5052	100.0	947	24	US-09-949-003C-2282
7	5052	100.0	947	25	US-09-981-397A-18
8	5052	100.0	947	29	US-10-394-322A-44
9	5052	100.0	947	37	US-60-366-892-44
10	5052	100.0	947	37	US-60-474-850-493
11	5040	99.8	947	26	US-10-087-192-888
12	5040	99.8	947	27	US-10-170-205B-8384
13	5040	99.8	947	27	US-10-170-205B-9184
14	5040	99.8	947	37	US-60-443-566-2590
15	5040	99.8	947	37	US-60-443-566-2591
16	5040	99.8	947	37	US-60-453-050-13077
17	5040	99.8	947	37	US-60-453-050-13078
18	5040	99.8	947	37	US-60-453-135-13077
19	5040	99.8	947	37	US-60-453-135-13078
20	5040	99.8	947	37	US-60-455-444-7013
21	5040	99.8	947	37	US-60-455-444-7014
22	5040	99.8	947	37	US-60-465-241-7013
23	5040	99.8	947	37	US-60-465-241-7014
24	5040	99.8	947	37	US-60-466-412-13077
25	5040	99.8	947	37	US-60-466-412-13078
26	5040	99.8	947	37	US-60-485-450-1471
27	5040	99.8	947	37	US-60-485-450-1472
28	5040	99.8	947	37	US-60-487-610-2325
29	5040	99.8	947	37	US-60-487-610-2326
30	5040	99.8	947	37	US-60-563-440-1471
31	5040	99.8	947	37	US-60-563-440-1472
32	5040	99.8	947	37	US-60-582-609-2325
33	5040	99.8	947	37	US-60-582-609-2326
34	5040	99.8	952	24	US-09-949-003C-3136
35	4946	97.9	939	37	US-60-248-798-348
36	4136.5	83.1	942	22	US-09-791-537-137083
37	3439.5	68.1	697	1	PCT-US03-26780-1811
38	2481.5	49.1	548	26	US-10-087-192-885
39	2301	45.5	477	33	US-10-760-320A-4878
40	2301	45.5	477	33	US-10-760-320A-4878
41	1572.5	31.1	368	22	US-09-758-457-758
42	1572.5	31.1	368	28	US-10-219-925-758
43	1028	20.3	188	37	US-60-195-052-734
44	992	19.6	184	37	US-60-196-711-1924
45	586	11.6	113	37	US-60-213-847-1073

ALIGNMENTS

RESULT 1  
US-09-155-676-7  
; Sequence 7, Application US/09155676  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH-21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 947 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-155-676-7

Query Match 100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLBVAVKSPVFCQKWLND 60
DB 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLBVAVKSPVFCQKWLND 60
QY 61 VITKGTAKGSEAGPAISIIAQAECNSQEFPTSERIFIAQSKQYSESILDIQINN 120
DB 61 VITKGTAKGSEAGPAISIIAQAECNSQEFPTSERIFIAQSKQYSESILDIQINN 120
QY 121 VAHATEGMAVCWKGRKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
DB 121 VAHATEGMAVCWKGRKRSKARKKRSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180
QY 181 DESPLGAPYVRNTQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
DB 181 DESPLGAPYVRNTQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPQDGGPLPLTHPPFPYSLRPLPFPFHPPLQPKPPLSFGLKACVDSQKPLDPDH 300
DB 241 KLHHPQDGGPLPLTHPPFPYSLRPLPFPFHPPLQPKPPLSFGLKACVDSQKPLDPDH 300
QY 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEYLVHVLQGVSSQASHLSLAKTW 360
DB 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEYLVHVLQGVSSQASHLSLAKTW 360
QY 361 AARGSRSPSKTDENEGVLITKLPVDYVEEVHWAHOLRLGRGSGFGEVHRMEDK 420
DB 361 AARGSRSPSKTDENEGVLITKLPVDYVEEVHWAHOLRLGRGSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLVFRABELMACAGLTSPTVPLYGAVREGPWNIWFMLLEGSLGQL 480
DB 421 QTGFQCAVKVRLVFRABELMACAGLTSPTVPLYGAVREGPWNIWFMLLEGSLGQL 480
QY 481 VKEQGCLPEDRALYIYGQALEGLYHLSRIILHGVKADNVLLSSDGGSHAALCDFGHAVC 540
DB 481 VKEQGCLPEDRALYIYGQALEGLYHLSRIILHGVKADNVLLSSDGGSHAALCDFGHAVC 540

RESULT 2
US-09-155-676-14
Sequence 14, Application US/09155676
GENERAL INFORMATION:
APPLICANT: WALLACH, David
APPLICANT: MALININ, Nikolai
APPLICANT: BOLDIN, Mark
APPLICANT: KOVALENKO, Andrei
APPLICANT: METT, Igor
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: BROWDY AND NEIMAR, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH-21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
```

```
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 947 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-155-676-14

Query Match      100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120

Qy 121 VAHATEGMARVCWKGRSKARKKXKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
Db 121 VAHATEGMARVCWKGRSKARKKXKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180

Qy 181 DESPLGAPYVNTPOFTKPLKEPGLGQLCFKQLGEGLRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVNTPOFTKPLKEPGLGQLCFKQLGEGLRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHPDQGGPLPLTHPPPPYRSLRPHPPFHPPLQPKWPHLSFLGKLCACVDSQKPLDPDH 300
Db 241 KLHPDQGGPLPLTHPPPPYRSLRPHPPFHPPLQPKWPHLSFLGKLCACVDSQKPLDPDH 300

Qy 301 LSKLACVDSKPLDGLPHLEPCLSRGAHEKESVEEYLVAHQGVSSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSKPLDGLPHLEPCLSRGAHEKESVEEYLVAHQGVSSVSSQAHSLSLAKTW 360

Qy 361 AARGSRSEPSPKTNEGVLITTEKLKPDVYEEVEVHWATHQLRGLRGSGFGEVHRMEDK 420
Db 361 AARGSRSEPSPKTNEGVLITTEKLKPDVYEEVEVHWATHQLRGLRGSGFGEVHRMEDK 420

Qy 421 QTGFQCAVKVRLEVFRABELMACAGLTSRIVPLYGAVREGPWVNIIPMELLEGGSLGQL 480
Db 421 QTGFQCAVKVRLEVFRABELMACAGLTSRIVPLYGAVREGPWVNIIPMELLEGGSLGQL 480

Qy 481 VKEQGCLPEDRALYYLGOALGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYYLGOALGLEYLHRSRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540

Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLHMLNGCHPWTQFF 600

Qy 601 RGPLCLKTASPPPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAELGGKVNRLAQVVG 660
Db 601 RGPLCLKTASPPPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAELGGKVNRLAQVVG 660

Qy 661 LKSPWRGKYKPRPPPPNQANYHTLHAQPRELSPRAPCPRAETTTGRAPKLPPLPPE 720
Db 661 LKSPWRGKYKPRPPPPNQANYHTLHAQPRELSPRAPCPRAETTTGRAPKLPPLPPE 720

Qy 721 PPEPNKSPPLTLKSEESGMWPEPLPSSLEPAPARNPSSPERKATVPPEOLOOLETEFLN 780
Db 721 PPEPNKSPPLTLKSEESGMWPEPLPSSLEPAPARNPSSPERKATVPPEOLOOLETEFLN 780

Qy 781 SLSQPFSLSEOEQILSLCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLSEOEQILSLCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSWSSQAEARSS 840

Qy 841 SWNVYLARGRTDTPSYFNGVKVQIQSLNGBHLHIREPHRVKVGDIATGSISSQIPAAAFS 900
Db 841 SWNVYLARGRTDTPSYFNGVKVQIQSLNGBHLHIREPHRVKVGDIATGSISSQIPAAAFS 900

Qy 901 LVTKDGQPVRYDMVDPDSDGLDQLTAPDGSFAMSWRVKQGLENRP 947
Db 901 LVTKDGQPVRYDMVDPDSDGLDQLTAPDGSFAMSWRVKQGLENRP 947

RESULT 3
US-09-155-676A-7
; Sequence 7, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155.676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IL97/00117
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25, 618
; REFERENCE/DOCKET NUMBER: WALLACH-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 947 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-155-676A-7

Query Match      100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMCPGAGSAGVQOKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAQAECENSQBFSPFTFSERIFIAGSKQYSQESLDQIPNN 120

Qy 121 VAHATEGMARVCWKGRSKARKKXKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
Db 121 VAHATEGMARVCWKGRSKARKKXKSKSLAHAGVALAKPLPTPEQESCTIPVOE 180
```

121 VAHATGKQAVRWKVKGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
181 DESPLGAPYVRNTPTQFTKPEGLGQLCFKQGLGRLPALPRSELHKLISPLQCLNHVW 240  
181 DESPLGAPYVRNTPTQFTKPEGLGQLCFKQGLGRLPALPRSELHKLISPLQCLNHVW 240  
241 KLHHPDQGGPLPLPHTPPYSLRPHFPHPLOPKPHPLSFGLKACVDSQKPLPDH 300  
241 KLHHPDQGGPLPLPHTPPYSLRPHFPHPLOPKPHPLSFGLKACVDSQKPLPDH 300  
301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360  
301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360  
361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHATHQLRGRSGFGEVHRMEDK 420  
361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHATHQLRGRSGFGEVHRMEDK 420  
421 QTGFQCAVKKVRLEVPRAELMACAGLTSRIVPLYGAVREGPWNI FMELEGGSLGQL 480  
421 QTGFQCAVKKVRLEVPRAELMACAGLTSRIVPLYGAVREGPWNI FMELEGGSLGQL 480  
481 VKEQGLCPEDRALYILGOALEGLEYLHSRRIILHGDVKNVLLSSDGSAAALCDFGHAVC 540  
481 VKEQGLCPEDRALYILGOALEGLEYLHSRRIILHGDVKNVLLSSDGSAAALCDFGHAVC 540  
541 LQPDGLGKSLLTGDIYPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQPF 600  
541 LQPDGLGKSLLTGDIYPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQPF 600  
601 RGPLCLKIASPPVREIPSCAPLTAQALQGLRKEPIHRYSAABELGKVNRAALQOVGG 660  
601 RGPLCLKIASPPVREIPSCAPLTAQALQGLRKEPIHRYSAABELGKVNRAALQOVGG 660  
661 LKSPWRGEYKEPRHPNPQANYHTLHAOPRELSPRAPGRPAEETTGAPKLQPLPPE 720  
661 LKSPWRGEYKEPRHPNPQANYHTLHAOPRELSPRAPGRPAEETTGAPKLQPLPPE 720  
721 PPEPNKSPPTLTKESGEMWEPLPLSLEPAPARNPSSPERKATVPEQLQLELFLN 780  
721 PPEPNKSPPTLTKESGEMWEPLPLSLEPAPARNPSSPERKATVPEQLQLELFLN 780  
781 SLSPFFLEQEQILSLCLSDLSLSDSEKNPSKASQSRDLSLSSGVHSSWSQAARS 840  
781 SLSPFFLEQEQILSLCLSDLSLSDSEKNPSKASQSRDLSLSSGVHSSWSQAARS 840  
841 SNNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFS 900  
841 SNNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFS 900  
901 LVTKDQGPVRYDMVPPSGIDLOCTLAPDGSFAWSRVRKHQLENRP 947  
901 LVTKDQGPVRYDMVPPSGIDLOCTLAPDGSFAWSRVRKHQLENRP 947

RESULT 4  
US-09-155-676A-14  
; Sequence 14, Application US/09155676A  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA

ZIP: 20001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/155,676A  
FILING DATE: 04-JAN-1999  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/IL97/00117  
FILING DATE: 01-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IL 117800  
FILING DATE: 02-APR-1996  
APPLICATION DATA:  
APPLICATION NUMBER: IL 119133  
FILING DATE: 26-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: WALLACH-21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 947 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-155-676A-14

Query Match 100.0%; Score 5052; DB 15; Length 947;  
Best Local Similarity 100.0%; Pred. No. 5.5e-306; Indels 0; Gaps 0;  
Matches 947; Conservative 0; Mismatches 0;  
QY 1 MAVMEMACGAPGAVGQOKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACGAPGAVGQOKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGKWEILND 60  
QY 61 VITKTAKEGSAGAPAAISIIAQACENSQSFPTFSERIIFAGSKQYSQESLDOIENN 120  
Db 61 VITKTAKEGSAGAPAAISIIAQACENSQSFPTFSERIIFAGSKQYSQESLDOIENN 120  
QY 121 VAHATGKQAVRWKVKGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
Db 121 VAHATGKQAVRWKVKGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
QY 181 DESPLGAPYVRNTPTQFTKPEGLGQLCFKQGLGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVRNTPTQFTKPEGLGQLCFKQGLGRLPALPRSELHKLISPLQCLNHVW 240  
QY 241 KLHHPDQGGPLPLPHTPPYSLRPHFPHPLOPKPHPLSFGLKACVDSQKPLPDH 300  
Db 241 KLHHPDQGGPLPLPHTPPYSLRPHFPHPLOPKPHPLSFGLKACVDSQKPLPDH 300  
QY 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360  
Db 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360  
QY 361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHATHQLRGRSGFGEVHRMEDK 420  
Db 361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHATHQLRGRSGFGEVHRMEDK 420  
QY 421 QTGFQCAVKKVRLEVPRAELMACAGLTSRIVPLYGAVREGPWNI FMELEGGSLGQL 480  
Db 421 QTGFQCAVKKVRLEVPRAELMACAGLTSRIVPLYGAVREGPWNI FMELEGGSLGQL 480  
QY 481 VKEQGLCPEDRALYILGOALEGLEYLHSRRIILHGDVKNVLLSSDGSAAALCDFGHAVC 540

Db 481 VKEQCLPEDRALYYLGOALEGLVLSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540  
Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Qy 601 RGPLCLKTASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRAALQQVGG 660  
Db 601 RGPLCLKTASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRAALQQVGG 660  
Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPAEETTCRAPKLOPPLPPE 720  
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPAEETTCRAPKLOPPLPPE 720  
Qy 721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQQLRIELFLN 780  
Db 721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQQLRIELFLN 780  
Qy 781 SLSQPFSLKEEQILCLSIDSLSDSEKNPKSKASOSSRDITLSSGVHSSQAARSS 840  
Db 781 SLSQPFSLKEEQILCLSIDSLSDSEKNPKSKASOSSRDITLSSGVHSSQAARSS 840  
Qy 841 SNNVLAGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900  
Db 841 SNNVLAGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900  
Qy 901 LVTKDGPVRYDMEVPSGIDLOCTLAPDGSFAMSWRVKHCQLENRP 947  
Db 901 LVTKDGPVRYDMEVPSGIDLOCTLAPDGSFAMSWRVKHCQLENRP 947

RESULT 5  
US-09-791-537-1606  
; Sequence 1606, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1606  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-1606

Query Match 100.0%; Score 5052; DB 22; Length 947;  
Best Local Similarity 100.0%; Pred. No. 5.5e-306;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60  
Qy 61 VITKGTAKESGAGPAALISIAQAECSQSFPTFSRIFRIFAGSKQYSQSESIDQIPNN 120  
Db 61 VITKGTAKESGAGPAALISIAQAECSQSFPTFSRIFRIFAGSKQYSQSESIDQIPNN 120  
Qy 121 VAHATEGMARVCKGKRRSKARKKRSKSLAHAGVALAKPLPRTPPEQESTIPVOE 180  
Db 121 VAHATEGMARVCKGKRRSKARKKRSKSLAHAGVALAKPLPRTPPEQESTIPVOE 180  
Qy 181 DESPLGAPYVNTPTQFTKPLKEPGLQGLCFKQGLGELRPAALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVNTPTQFTKPLKEPGLQGLCFKQGLGELRPAALPRSELHKLISPLQCLNHVW 240  
Qy 241 KLHPDQGGPLPLPTHPPPPYRRLPHPPFHPQWKPHPLESFLGKLCACVDSQKPLDPDPH 300

Db 241 KLHPDQGGPLPLPTHPPPPYRRLPHPPFHPQWKPHPLESFLGKLCACVDSQKPLDPDPH 300  
Qy 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVLHALQGSVSSQAHSLSLAKTW 360  
Db 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVLHALQGSVSSQAHSLSLAKTW 360  
Qy 361 AARGSRSEPSKPTEDNEGVLITTEKLPVDYEEVEVHWATHQLRLGRSGFGEVHRMEDK 420  
Db 361 AARGSRSEPSKPTEDNEGVLITTEKLPVDYEEVEVHWATHQLRLGRSGFGEVHRMEDK 420  
Qy 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLVYGAVERGPWWNIEMELLEGSLQGL 480  
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLVYGAVERGPWWNIEMELLEGSLQGL 480  
Qy 481 VKEQCLPEDRALYYLGOALEGLVLSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540  
Db 481 VKEQCLPEDRALYYLGOALEGLVLSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540  
Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Qy 601 RGPLCLKTASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRAALQQVGG 660  
Db 601 RGPLCLKTASBPVPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRAALQQVGG 660  
Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPAEETTCRAPKLOPPLPPE 720  
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPPAEETTCRAPKLOPPLPPE 720  
Qy 721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQQLRIELFLN 780  
Db 721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQQLRIELFLN 780  
Qy 781 SLSQPFSLKEEQILCLSIDSLSDSEKNPKSKASOSSRDITLSSGVHSSQAARSS 840  
Db 781 SLSQPFSLKEEQILCLSIDSLSDSEKNPKSKASOSSRDITLSSGVHSSQAARSS 840  
Qy 841 SNNVLAGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900  
Db 841 SNNVLAGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900  
Qy 901 LVTKDGPVRYDMEVPSGIDLOCTLAPDGSFAMSWRVKHCQLENRP 947  
Db 901 LVTKDGPVRYDMEVPSGIDLOCTLAPDGSFAMSWRVKHCQLENRP 947

RESULT 6  
US-09-949-003C-2282  
; Sequence 2282, Application US/09949003C  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH NEUROPSYCHIATRIC  
; TITLE OF INVENTION: DISORDERS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL000791  
; CURRENT APPLICATION NUMBER: US/09/949,003C  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/231,446  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 74065  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 2282  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-949-003C-2282

Query Match 100.0%; Score 5052; DB 24; Length 947;  
Best Local Similarity 100.0%; Pred. No. 5.5e-306;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

```
Db 1 MAVMEMACPGAGSAGVQOKELPKPKETPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAAECENSQESPTFSERIFTAGSKQYSQSSLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAAECENSQESPTFSERIFTAGSKQYSQSSLDQIPNN 120
QY 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLRTPTEQESCTIPVOE 180
Db 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLRTPTEQESCTIPVOE 180
QY 181 DESPLGAPVVRNTPOFTKPLPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPOFTKPLPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPDQGGPLPLPTHPPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 300
Db 241 KLHHPDQGGPLPLPTHPPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 300
QY 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
QY 361 AARGSRSPSPKTEDENEGVLITTEKLKPDVYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AARGSRSPSPKTEDENEGVLITTEKLKPDVYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGOL 480
QY 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMMLHMLNGCHPWTQPF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMMLHMLNGCHPWTQPF 600
QY 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQQVGG 660
Db 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQQVGG 660
QY 661 LKSPMRGEYKEPRHPPNQANYHOTLHAQPRELSRAPRPAEETTGRAPKLQPLPPE 720
Db 661 LKSPMRGEYKEPRHPPNQANYHOTLHAQPRELSRAPRPAEETTGRAPKLQPLPPE 720
QY 721 PPEPNKSPPLTSLKESGMEWELPLSLSEAPARNPSSPERKATVPEQLOQLELFLN 780
Db 721 PPEPNKSPPLTSLKESGMEWELPLSLSEAPARNPSSPERKATVPEQLOQLELFLN 780
QY 781 SLSPQFSLEREOQILSLCLSIDSLSDSEKKNPSKASQSSRDLTSSGVHWSWSQAEARS 840
Db 781 SLSPQFSLEREOQILSLCLSIDSLSDSEKKNPSKASQSSRDLTSSGVHWSWSQAEARS 840

; TITLE OF INVENTION: Cytomegalovirus Infection and their
; Sequence 18, Application US/09981397A-18
; GENERAL INFORMATION:
; APPLICANT: Axxima Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberger, Peter
; APPLICANT: Stein-Gerlach, Mathias
; APPLICANT: Bevec, Dorian
; TITLE OF INVENTION: Cytomegalovirus Infection and their
```

```
; TITLE OF INVENTION: Inhibition
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981,397A
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-981-397A-18

Query Match 100.0%; Score 5052; DB 25; Length 947;
Best Local Similarity 100.0%; Pred. No. 5.5e-306;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQOKELPKPKETPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQOKELPKPKETPLPGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGPAAISIIAAECENSQESPTFSERIFTAGSKQYSQSSLDQIPNN 120
Db 61 VITKGTAKGSEAGPAAISIIAAECENSQESPTFSERIFTAGSKQYSQSSLDQIPNN 120
QY 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLRTPTEQESCTIPVOE 180
Db 121 VAHATEGKMARVCKWKGKRSKARKKXKSSSLAHAGVALAKPLRTPTEQESCTIPVOE 180
QY 181 DESPLGAPVVRNTPOFTKPLPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVVRNTPOFTKPLPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPDQGGPLPLPTHPPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 300
Db 241 KLHHPDQGGPLPLPTHPPFPYSLRPLPGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 300
QY 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSSQAHSLTSLAKTW 360
QY 361 AARGSRSPSPKTEDENEGVLITTEKLKPDVYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
Db 361 AARGSRSPSPKTEDENEGVLITTEKLKPDVYREVEVHWATHQLRLGRSGFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGOL 480
QY 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYILGQALEGLEYLHRSRRIILHGDVKADNVLLSDGSHAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMMLHMLNGCHPWTQPF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRCDKADVWSSCCMMLHMLNGCHPWTQPF 600
QY 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQQVGG 660
Db 601 RGPLCLKIASBPVPVREIIPSCAPLTAQAIQGLRKEPIHRYSAALGKGNRALQQVGG 660
QY 661 LKSPMRGEYKEPRHPPNQANYHOTLHAQPRELSRAPRPAEETTGRAPKLQPLPPE 720
Db 661 LKSPMRGEYKEPRHPPNQANYHOTLHAQPRELSRAPRPAEETTGRAPKLQPLPPE 720
QY 721 PPEPNKSPPLTSLKESGMEWELPLSLSEAPARNPSSPERKATVPEQLOQLELFLN 780
Db 721 PPEPNKSPPLTSLKESGMEWELPLSLSEAPARNPSSPERKATVPEQLOQLELFLN 780
QY 781 SLSPQFSLEREOQILSLCLSIDSLSDSEKKNPSKASQSSRDLTSSGVHWSWSQAEARS 840
Db 781 SLSPQFSLEREOQILSLCLSIDSLSDSEKKNPSKASQSSRDLTSSGVHWSWSQAEARS 840
```

Qy 841 SNNVYLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Db 841 SNNVYLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Qy 901 LVTKDGOVRYDMEVDPGSDGLDQCTLAPDGSFAWSRWKVGQLENRP 947  
Db 901 LVTKDGOVRYDMEVDPGSDGLDQCTLAPDGSFAWSRWKVGQLENRP 947  
RESULT 8  
US-10-394-322A-44  
; Sequence 44, Application US/10394322A  
; GENERAL INFORMATION:  
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: 39750-0006 US  
; CURRENT APPLICATION NUMBER: US/10/394,322A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: US 60/366,892  
; PRIOR FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-394-322A-44  
Query Match 100.0%; Score 5052; DB 29; Length 947;  
Best Local Similarity 100.0%; Pred. No. 5.5e-306;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120  
Db 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120  
Qy 121 VAHATEGMARVCKWKGRSKARKKRRKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
Db 121 VAHATEGMARVCKWKGRSKARKKRRKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
Qy 181 DESPLGAPYVRNTPTFTKPLKEPGLQCFKQKGEGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVRNTPTFTKPLKEPGLQCFKQKGEGRLPALPRSELHKLISPLQCLNHVW 240  
Qy 241 KLHPDQGGPLPLPHTPPYRLPFPFHPPLQPKWHPHLESFLGKLACVDSQKPLDPDH 300  
Db 241 KLHPDQGGPLPLPHTPPYRLPFPFHPPLQPKWHPHLESFLGKLACVDSQKPLDPDH 300  
Qy 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360  
Db 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360  
Qy 361 AARGSRREPSKPTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRGSFGEVHRMEDK 420  
Db 361 AARGSRREPSKPTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRGSFGEVHRMEDK 420  
Qy 421 QTGFOCAVKVRLEVFRAEELMACGLTSPRIVPLYGAVREGPWNIIFMELLEGSLGQL 480  
Db 421 QTGFOCAVKVRLEVFRAEELMACGLTSPRIVPLYGAVREGPWNIIFMELLEGSLGQL 480  
Qy 481 VKEQGLPEDRALYYLGALSGLEYLHSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540  
Db 481 VKEQGLPEDRALYYLGALSGLEYLHSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540  
Qy 541 LQPDGLGKSLTGDIYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLTGDIYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600

Qy 601 RGPLCLKIASPPVREIPPSCAPLTAQAIQEGRLKEPIHRVSAEELGGKVNRLAQVGG 660  
Db 601 RGPLCLKIASPPVREIPPSCAPLTAQAIQEGRLKEPIHRVSAEELGGKVNRLAQVGG 660  
Qy 661 LKSPWGEYKEPRHPPPNQANYHOTLHAQPRELSPRAGPRPAEETTGRAKLPPLPPE 720  
Db 661 LKSPWGEYKEPRHPPPNQANYHOTLHAQPRELSPRAGPRPAEETTGRAKLPPLPPE 720  
Qy 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSPERKATVPPEQLQOLELEFLN 780  
Db 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSPERKATVPPEQLQOLELEFLN 780  
Qy 781 SLSQPFSLERQEQILSLCLSDSLSDSEKNPKSKSQSSRDTLSSGVHSSWSSQAERSS 840  
Db 781 SLSQPFSLERQEQILSLCLSDSLSDSEKNPKSKSQSSRDTLSSGVHSSWSSQAERSS 840  
Qy 841 SNNVYLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Db 841 SNNVYLARGRPDTPTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Qy 901 LVTKDGOVRYDMEVDPGSDGLDQCTLAPDGSFAWSRWKVGQLENRP 947  
Db 901 LVTKDGOVRYDMEVDPGSDGLDQCTLAPDGSFAWSRWKVGQLENRP 947  
RESULT 9  
US-60-366-892-44  
; Sequence 44, Application US/60366892  
; GENERAL INFORMATION:  
; APPLICANT: Prescott, John C.  
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
; FILE REFERENCE: SUNESIS.006PR  
; CURRENT APPLICATION NUMBER: US/60/366,892  
; CURRENT FILING DATE: 2002-03-21  
; NUMBER OF SEQ ID NOS: 69  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-366-892-44  
Query Match 100.0%; Score 5052; DB 37; Length 947;  
Best Local Similarity 100.0%; Pred. No. 5.5e-306;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACPGAGSAGVQKQELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120  
Db 61 VITKGTAKGSEAGPAAISIIAQACENSQFSTFSERIFIAGSKQYSQESLDOIINN 120  
Qy 121 VAHATEGMARVCKWKGRSKARKKRRKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
Db 121 VAHATEGMARVCKWKGRSKARKKRRKSKSLAHAGVALAKPLPRTPEQESCTIPVQE 180  
Qy 181 DESPLGAPYVRNTPTFTKPLKEPGLQCFKQKGEGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVRNTPTFTKPLKEPGLQCFKQKGEGRLPALPRSELHKLISPLQCLNHVW 240  
Qy 241 KLHPDQGGPLPLPHTPPYRLPFPFHPPLQPKWHPHLESFLGKLACVDSQKPLDPDH 300  
Db 241 KLHPDQGGPLPLPHTPPYRLPFPFHPPLQPKWHPHLESFLGKLACVDSQKPLDPDH 300  
Qy 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360  
Db 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSLSLAKTW 360  
Qy 361 AARGSRREPSKPTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRGSFGEVHRMEDK 420  
Db 361 AARGSRREPSKPTEDNEGVLLTEKLPVDYVEEYEEVHWATHQLRLGRGSFGEVHRMEDK 420

Db 361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHWATHQLRGRSGFGEVHRMEDK 420  
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480  
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480  
QY 481 VKEQGCLPEDRALYILGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540  
Db 481 VKEQGCLPEDRALYILGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540  
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
QY 601 RGPLCLKIASEPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOVGG 660  
Db 601 RGPLCLKIASEPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOVGG 660  
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPGPRPAEETTGRAPKLQPPLPPE 720  
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPGPRPAEETTGRAPKLQPPLPPE 720  
QY 721 PPEPNKSPPLTUSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELOLEIELEFLN 780  
Db 721 PPEPNKSPPLTUSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELOLEIELEFLN 780  
QY 781 SLSQPFSLEEQEQLSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840  
Db 781 SLSQPFSLEEQEQLSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840  
QY 841 SNMVLARGRPDTPTPSYFNGVKVQIOSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Db 841 SNMVLARGRPDTPTPSYFNGVKVQIOSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
QY 901 LVTXGQOPVRYDMEVPSDGDIDLOCTLAPDGSFAMSWRVKHGOLENRP 947  
Db 901 LVTXGQOPVRYDMEVPSDGDIDLOCTLAPDGSFAMSWRVKHGOLENRP 947

RESULT 10  
US-60-474-850-493  
; Sequence 493, Application US/60474850  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Ly, Ngoc  
; APPLICANT: Prentice, James  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING  
; TITLE OF INVENTION: RENAL TRANSPLANT REJECTION  
; FILE REFERENCE: 506613000700  
; CURRENT APPLICATION NUMBER: US/60/474,850  
; CURRENT FILING DATE: 2003-05-30  
; NUMBER OF SEQ ID NOS: 552  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 493  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-474-850-493

Query Match 100.0%; Score 5052; DB 37; Length 947;  
Best Local Similarity 100.0%; Pred. No. 5.5e-306;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACGAGSAGVQOKELPKKEKTPPLGKKOSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACGAGSAGVQOKELPKKEKTPPLGKKOSSVYKLEAVEKSPVFCGKWEILND 60  
QY 61 VITKGTAKESGAGPAASIIIAQACENSQEFSPTFSEIFRIFAGSKQYSQSESLDQIPNN 120  
Db 61 VITKGTAKESGAGPAASIIIAQACENSQEFSPTFSEIFRIFAGSKQYSQSESLDQIPNN 120  
QY 121 VAHATEGKMARVCMWKGKRRSKARKKKSKSLAHAGVALAKPLPRTPEQESCTIPQOE 180

Db 121 VAHATEGKMARVCMWKGKRRSKARKKKSKSLAHAGVALAKPLPRTPEQESCTIPQOE 180  
QY 181 DESPLGAPYVRNTPOFTPLKEPGLGOLCFKQOLGEGLRPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVRNTPOFTPLKEPGLGOLCFKQOLGEGLRPALPRSELHKLISPLQCLNHVW 240  
QY 241 KLHHPQDGGPILPLTHPPFPYSRLEPFPHPPLQPKWPKPLESFLGKLACVDSOKPLPDPH 300  
Db 241 KLHHPQDGGPILPLTHPPFPYSRLEPFPHPPLQPKWPKPLESFLGKLACVDSOKPLPDPH 300  
QY 301 LSKLACVDSKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360  
Db 301 LSKLACVDSKPLPGLPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360  
QY 361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHWATHQLRGRSGFGEVHRMEDK 420  
Db 361 AARGSRSEPSPKTNEGVLLTEKLPVDYREVEVHWATHQLRGRSGFGEVHRMEDK 420  
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480  
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNI FWELEGGSLGOL 480  
QY 481 VKEQGCLPEDRALYILGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540  
Db 481 VKEQGCLPEDRALYILGQALEGLEYLEYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540  
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
QY 601 RGPLCLKIASEPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOVGG 660  
Db 601 RGPLCLKIASEPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGKGKVNRAALQOVGG 660  
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPGPRPAEETTGRAPKLQPPLPPE 720  
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPPAPGPRPAEETTGRAPKLQPPLPPE 720  
QY 721 PPEPNKSPPLTUSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELOLEIELEFLN 780  
Db 721 PPEPNKSPPLTUSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPQELOLEIELEFLN 780  
QY 781 SLSQPFSLEEQEQLSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840  
Db 781 SLSQPFSLEEQEQLSCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSSQAEARSS 840  
QY 841 SNMVLARGRPDTPTPSYFNGVKVQIOSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Db 841 SNMVLARGRPDTPTPSYFNGVKVQIOSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
QY 901 LVTXGQOPVRYDMEVPSDGDIDLOCTLAPDGSFAMSWRVKHGOLENRP 947  
Db 901 LVTXGQOPVRYDMEVPSDGDIDLOCTLAPDGSFAMSWRVKHGOLENRP 947

RESULT 11  
US-10-087-192-888  
; Sequence 888, Application US/10087192  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR FILING DATE: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR FILING DATE: US 09/798,586  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 888



; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-087-192-888

```
Query Match      99.8%; Score 5040; DB 26; Length 947;
Best Local Similarity 99.8%; Pred. No. 3.1e-305;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLPGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLPGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Qy 61 VITKGTAKGSEAGAPAAISIIAQAECSQSFPTFSERIFIAGSKQYSQSESIDQIPNN 120
Db 61 VITKGTAKGSEAGAPAAISIIAQAECSQSFPTFSERIFIAGSKQYSQSESIDQIPNN 120
Qy 121 VAHATEGKMARVCWKGRSKARKKKSKSLAHAGVALAKPLPRTPQEESCITIPVOE 180
Db 121 VAHATEGKMARVCWKGRSKARKKKSKSLAHAGVALAKPLPRTPQEESCITIPVOE 180
Qy 181 DESPLGAPYVNTPQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVNTPQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Qy 241 KLHHPDGGPLPLPHTPPYGRPLPHPPFHPPLQPKWPHLESFLGKLACVDSQKPLDPDH 300
Db 241 KLHHPDGGPLPLPHTPPYGRPLPHPPFHPPLQPKWPHLESFLGKLACVDSQKPLDPDH 300
Qy 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKSVSEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKSVSEYLVHALQGSVSSQAHSLSLAKTW 360
Qy 361 AARGSRSEPSPKTEDNEGVLITKLPVDYEEVEVHWATHQLRLGRGSFGEVHRMEDK 420
Db 361 AARGSRSEPSPKTEDNEGVLITKLPVDYEEVEVHWATHQLRLGRGSFGEVHRMEDK 420
Qy 421 QTGFQCAVKVVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGSLGQL 480
Db 421 QTGFQCAVKVVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGSLGQL 480
Qy 481 VKEQCCLPEDRALYYLGQALSGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQCCLPEDRALYYLGQALSGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASPPPPVREIPPSCAPLTAQAIQEGRLKEPIHRSVAELGGKVNRAALQVGG 660
Db 601 RGPLCLKIASPPPPVREIPPSCAPLTAQAIQEGRLKEPIHRSVAELGGKVNRAALQVGG 660
Qy 661 LKSPWRGKYKEPRHPPPNQANYHOTLHAQRELSPRAPGPABETTGRAKLPPLPPE 720
Db 661 LKSPWRGKYKEPRHPPPNQANYHOTLHAQRELSPRAPGPABETTGRAKLPPLPPE 720
Qy 721 PPEPKNSPPLTLSKEESGMWEPPLSLSEAPARNPSSPERKATVPEQELQQLIEFLN 780
Db 721 PPEPKNSPPLTLSKEESGMWEPPLSLSEAPARNPSSPERKATVPEQELQQLIEFLN 780
Qy 781 SLSQPFSLSEQEQILSLISLSDSEKNPKSKASSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLSEQEQILSLISLSDSEKNPKSKASSRDTLSSGVHSSWSSQAEARSS 840
Qy 841 SNNMVLARGRTDTPSYNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISQIIPAAAFS 900
Db 841 SNNMVLARGRTDTPSYNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISQIIPAAAFS 900
Qy 901 LVTKDQVRYDMEVPDSDGIDLQCTLAPDGSFANSWRVKGQLENRP 947
Db 901 LVTKDQVRYDMEVPDSDGIDLQCTLAPDGSFANSWRVKGQLENRP 947
```

RESULT 12  
US-10-170-205E-8384

```
; Sequence 8384, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE REFERENCE: CLO01381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8384
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-8384
```

```
Query Match      99.8%; Score 5040; DB 27; Length 947;
Best Local Similarity 99.8%; Pred. No. 3.1e-305;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLPGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLPGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Qy 61 VITKGTAKGSEAGAPAAISIIAQAECSQSFPTFSERIFIAGSKQYSQSESIDQIPNN 120
Db 61 VITKGTAKGSEAGAPAAISIIAQAECSQSFPTFSERIFIAGSKQYSQSESIDQIPNN 120
Qy 121 VAHATEGKMARVCWKGRSKARKKKSKSLAHAGVALAKPLPRTPQEESCITIPVOE 180
Db 121 VAHATEGKMARVCWKGRSKARKKKSKSLAHAGVALAKPLPRTPQEESCITIPVOE 180
Qy 181 DESPLGAPYVNTPQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVNTPQFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Qy 241 KLHHPDGGPLPLPHTPPYGRPLPHPPFHPPLQPKWPHLESFLGKLACVDSQKPLDPDH 300
Db 241 KLHHPDGGPLPLPHTPPYGRPLPHPPFHPPLQPKWPHLESFLGKLACVDSQKPLDPDH 300
Qy 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKSVSEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKSVSEYLVHALQGSVSSQAHSLSLAKTW 360
Qy 361 AARGSRSEPSPKTEDNEGVLITKLPVDYEEVEVHWATHQLRLGRGSFGEVHRMEDK 420
Db 361 AARGSRSEPSPKTEDNEGVLITKLPVDYEEVEVHWATHQLRLGRGSFGEVHRMEDK 420
Qy 421 QTGFQCAVKVVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGSLGQL 480
Db 421 QTGFQCAVKVVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGSLGQL 480
Qy 481 VKEQCCLPEDRALYYLGQALSGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQCCLPEDRALYYLGQALSGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASPPPPVREIPPSCAPLTAQAIQEGRLKEPIHRSVAELGGKVNRAALQVGG 660
Db 601 RGPLCLKIASPPPPVREIPPSCAPLTAQAIQEGRLKEPIHRSVAELGGKVNRAALQVGG 660
Qy 661 LKSPWRGKYKEPRHPPPNQANYHOTLHAQRELSPRAPGPABETTGRAKLPPLPPE 720
Db 661 LKSPWRGKYKEPRHPPPNQANYHOTLHAQRELSPRAPGPABETTGRAKLPPLPPE 720
Qy 721 PPEPKNSPPLTLSKEESGMWEPPLSLSEAPARNPSSPERKATVPEQELQQLIEFLN 780
Db 721 PPEPKNSPPLTLSKEESGMWEPPLSLSEAPARNPSSPERKATVPEQELQQLIEFLN 780
```

Db 721 PPEPNKSPPLTLTSLKEESGMWEPLPLSLLEPAPAPARNPSSPERKATVPEQELQLELFLN 780  
QY 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840  
Db 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840  
QY 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDVGIATGSISSQIPAAAFS 900  
Db 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDVGIATGSISSQIPAAAFS 900  
QY 901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSRVVKHGQLENRP 947  
Db 901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSRVVKHGQLENRP 947  
  
RESULT 13  
US-10-170-205E-9184  
; Sequence 9184, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 9184  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-9184

Query Match 99.8%; Score 5040; DB 27; Length 947;  
Best Local Similarity 99.8%; Pred. No. 3.1e-305;  
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 MAVMEMACFGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACFGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60  
  
QY 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFTAGSKQYSQSSLSQIPNN 120  
Db 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFTAGSKQYSQSSLSQIPNN 120  
  
QY 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180  
Db 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180  
  
QY 181 DESPLGAPVVRNTPQTKLPKPEGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPVVRNTPQTKLPKPEGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240  
  
QY 241 KLHHPDQGGPLPLPHTPPFYSRLPHFPFPHLPQPKPHEPFLGKLCACVDSQKPLPDPH 300  
Db 241 KLHHPDQGGPLPLPHTPPFYSRLPHFPFPHLPQPKPHEPFLGKLCACVDSQKPLPDPH 300  
  
QY 301 LSKLACVDSQKPLPDGPHLEPSCLSRGAHEKFSVEEYLVAHALQGSVSSGQAHSLSLAKTW 360  
Db 301 LSKLACVDSQKPLPDGPHLEPSCLSRGAHEKFSVEEYLVAHALQGSVSSGQAHSLSLAKTW 360  
  
QY 361 AARGSRSPSKPTEDNEGVLITKLPVDYREEVHWAHQRLGRGSFGEVHRMEDK 420  
Db 361 AARGSRSPSKPTEDNEGVLITKLPVDYREEVHWAHQRLGRGSFGEVHRMEDK 420  
  
QY 421 QTGFQCAVKKVRLEVFRAELMACAGLTSPRIVPLVYCAVREGPWNVIFMELLEGGSLGQL 480  
Db 421 QTGFQCAVKKVRLEVFRAELMACAGLTSPRIVPLVYCAVREGPWNVIFMELLEGGSLGQL 480  
  
QY 481 VKEQGLCPEDRALYILGOALEGLEYLHSRRIILHGDVKADNVLLSSDGSAAALCDFGHAYC 540  
Db 481 VKEQGLCPEDRALYILGOALEGLEYLHSRRIILHGDVKADNVLLSSDGSAAALCDFGHAYC 540

QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600  
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600  
  
QY 601 RGPLCLKIASPPPPVREIIPSCAPLTAQAIQEGLRKEPIHVRVSAELGKKNRALKQOQVG 660  
Db 601 RGPLCLKIASPPPPVREIIPSCAPLTAQAIQEGLRKEPIHVRVSAELGKKNRALKQOQVG 660  
  
QY 661 LKSPMRGKYKPRHPPPNOANYHOTLHAQPRELSPRAPGPRPAEETTCRAPKLOPPLPPE 720  
Db 661 LKSPMRGKYKPRHPPPNOANYHOTLHAQPRELSPRAPGPRPAEETTCRAPKLOPPLPPE 720  
  
QY 721 PPEPNKSPPLTLTSLKEESGMWEPLPLSLLEPAPAPARNPSSPERKATVPEQELQLELFLN 780  
Db 721 PPEPNKSPPLTLTSLKEESGMWEPLPLSLLEPAPAPARNPSSPERKATVPEQELQLELFLN 780  
  
QY 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840  
Db 781 SLSQPFSLBEQIILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSSQAEARSS 840  
  
QY 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDVGIATGSISSQIPAAAFS 900  
Db 841 SNNMVLARGRPDTPDSYFNGVKVQIQSLNGEHLHIREFHRVKVGDVGIATGSISSQIPAAAFS 900  
  
QY 901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSRVVKHGQLENRP 947  
Db 901 LVTKDQGPVRYDMEVPDSDGIDLQCTLAPDGSFAMSRVVKHGQLENRP 947  
  
RESULT 14  
US-60-443-566-2590  
; Sequence 2590, Application US/60443566  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: BEGOVICH, Ann  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001447  
; CURRENT APPLICATION NUMBER: US/60/443,566  
; CURRENT FILING DATE: 2003-01-30  
; NUMBER OF SEQ ID NOS: 25102  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2590  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-443-566-2590

Query Match 99.8%; Score 5040; DB 37; Length 947;  
Best Local Similarity 99.8%; Pred. No. 3.1e-305;  
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 MAVMEMACFGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACFGAGSAGVQKQKELPKPKETPPGKQSSVYKLEAVEKSPVFCGKWEILND 60  
  
QY 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFTAGSKQYSQSSLSQIPNN 120  
Db 61 VITKTAGEGSBAGPAAISIIAQAECENSQEPSPTFSERIFTAGSKQYSQSSLSQIPNN 120  
  
QY 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180  
Db 121 VAHATEGKMARVCWKGRSKARKKRRKKSSKSLAHAGVALAKPLPRTPEQSSCTIPVQE 180  
  
QY 181 DESPLGAPVVRNTPQTKLPKPEGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPVVRNTPQTKLPKPEGLQGLCFKOLGEGRLPALPRSELHKLISPLQCLNHVW 240  
  
QY 241 KLHHPDQGGPLPLPHTPPFYSRLPHFPFPHLPQPKPHEPFLGKLCACVDSQKPLPDPH 300  
Db 241 KLHHPDQGGPLPLPHTPPFYSRLPHFPFPHLPQPKPHEPFLGKLCACVDSQKPLPDPH 300  
  
QY 301 LSKLACVDSQKPLPDGPHLEPSCLSRGAHEKFSVEEYLVAHALQGSVSSGQAHSLSLAKTW 360

```
Db 301 LSKLACVDSKPPLPGPHLEPSCLSGAHEKSVSEYLVHALQGSVSSQAHSLTSLAKTW 360
Qy 361 AARGSRSEPPKPTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRLGRGSGFGEVHRMEDK 420
Db 361 AARGSRSEPPKPTEDNEGVLITTEKLPVDYVEEVEVHWATHQLRLGRGSGFGEVHRMEDK 420
Qy 421 QTGFQCAVKKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGOL 480
Db 421 QTGFQCAVKKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGOL 480
Qy 481 VKEQCCLPEDRALYVLGOALEGLVYHSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540
Db 481 VKEQCCLPEDRALYVLGOALEGLVYHSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540
Qy 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600
Qy 601 RGPLCLKIASPPPPVREIPPPSCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRLAQVGG 660
Db 601 RGPLCLKIASPPPPVREIPPPSCAPLTAQAIQEGLRKEPIHRVSAELGGKVNRLAQVGG 660
Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSAPRAPGPABETTTGRAPKLOPPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSAPRAPGPABETTTGRAPKLOPPLPPE 720
Qy 721 PPEPNKSPPLTSLKEESGMWEPPLSLLEPAPARNPSSPERKATVPPEQLQOOLETELFLN 780
Db 721 PPEPNKSPPLTSLKEESGMWEPPLSLLEPAPARNPSSPERKATVPPEQLQOOLETELFLN 780
Qy 781 SLSQPFSLSEEQILSLCLSIDSLSDSDSEKNPSKASQSSRDITLSSGVHSSSQAEARSS 840
Db 781 SLSQPFSLSEEQILSLCLSIDSLSDSDSEKNPSKASQSSRDITLSSGVHSSSQAEARSS 840
Qy 841 SNNVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFS 900
Db 841 SNNVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFS 900
Qy 901 LVTKDQGPVRYDMEVPDSDGLDQCTLAPDGSFAMSWRVKXGQLENRP 947
Db 901 LVTKDQGPVRYDMEVPDSDGLDQCTLAPDGSFAMSWRVKXGQLENRP 947

RESULT 15
US-60-443-566-2591
; Sequence 2591, Application US/60443566
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001447
; CURRENT APPLICATION NUMBER: US/60/443.566
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 25102
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2591
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-443-566-2591

Query Match 99.8%; Score 5040; DB 37; Length 947;
Best Local Similarity 99.8%; Pred.No. 3.le-305;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQOKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQOKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60
Qy 61 VITGTAKGSEAGPAAISIIAQAECENSQBFSTFTSERIFIAGSKQYSQESLDQIPNN 120
```

Search completed: June 7, 2005, 12:19:39  
Job time : 427.493 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:51:20 ; Search time 37.0523 Seconds  
(without alignments)  
1907.914 Million cell updates/sec

Title: US-09-155-676B-7  
Perfect score: 5052  
Sequence: 1 MAVMEMACPGAGSGAVGQOK.....PDGSFAMSWRVKHQLENRP 947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pcp:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pcp:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pcp:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pcp:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pcp:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5052	100.0	947	3	US-09-257-703-1
2	5052	100.0	947	4	US-09-871-889A-1
3	5044	99.8	947	2	US-08-887-518-2
4	5044	99.8	947	2	US-09-023-321-2
5	5044	99.8	947	2	US-09-032-475-2
6	479	9.5	94	4	US-09-513-999C-6660
7	360.5	7.1	467	4	US-09-522-775A-4
8	360	7.1	467	4	US-09-522-775A-2
9	360	7.1	475	4	US-09-949-016-10456
10	358.5	7.1	1302	3	US-09-423-890-2
11	355.5	7.0	672	1	US-08-049-254-2
12	355.5	7.0	672	1	US-08-472-934-2
13	355.5	7.0	672	2	US-08-323-460A-2
14	355.5	7.0	672	2	US-08-461-146C-2
15	355.5	7.0	672	3	US-08-461-145C-2
16	355.5	7.0	672	3	US-08-628-829-2
17	355.5	7.0	1492	4	US-09-697-898-5
18	355.5	7.0	1493	3	US-09-423-890-8
19	355.5	7.0	1593	3	US-08-628-829-4
20	355	7.0	1512	4	US-09-697-898-2
21	353	7.0	619	1	US-08-472-934-4
22	353	7.0	619	2	US-08-323-460A-4
23	353	7.0	619	2	US-08-461-146C-4
24	353	7.0	619	3	US-08-461-145C-4
25	353	7.0	619	3	US-09-423-890-10
26	353	7.0	619	3	US-08-628-829-6
27	353	7.0	1492	4	US-09-697-898-4

28	352	7.0	619	1	US-08-472-934-12	Sequence 12, Appl
29	352	7.0	619	2	US-08-461-146C-12	Sequence 12, Appl
30	352	7.0	619	3	US-08-461-145C-12	Sequence 12, Appl
31	352	7.0	619	3	US-08-628-829-8	Sequence 8, Appl
32	348	6.9	619	3	US-09-423-890-4	Sequence 4, Appl
33	346	6.8	1495	4	US-09-697-898-3	Sequence 3, Appl
34	326.5	6.5	651	4	US-09-371-338-15	Sequence 15, Appl
35	322	6.4	268	4	US-09-371-338-17	Sequence 17, Appl
36	322	6.4	647	3	US-09-031-563-7	Sequence 7, Appl
37	322	6.4	647	4	US-09-392-277-7	Sequence 7, Appl
38	322	6.4	647	4	US-09-258-000-7	Sequence 7, Appl
39	322	6.4	648	3	US-09-031-563-5	Sequence 5, Appl
40	322	6.4	648	4	US-09-392-277-5	Sequence 5, Appl
41	322	6.4	648	4	US-09-258-000-5	Sequence 5, Appl
42	322	6.4	1315	3	US-09-031-563-2	Sequence 2, Appl
43	322	6.4	1315	3	US-09-293-505-10	Sequence 10, Appl
44	322	6.4	1315	4	US-09-392-277-2	Sequence 2, Appl
45	322	6.4	1315	4	US-09-258-000-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-257-703-1

; Sequence 1, Application US/09257703

; Patent No. 6265538

; GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

; APPLICANT: Gelezuinas, Romas

; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED

; FILE REFERENCE: 30448.61USU1

; CURRENT APPLICATION NUMBER: US/09/257,703

; CURRENT FILING DATE: 1999-02-25

; EARLIER APPLICATION NUMBER: 60/076,299

; EARLIER FILING DATE: 1998-02-27

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 947

; TYPE: PRT

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-257-703-1

Query Match	100.0%	Score	5052;	DB	3;	Length	947;
Best Local Similarity	100.0%;	Pred. No.	0;	Mismatches	0;	Indels	0;
Matches	947;	Conservative	0;				
Qy	1	MAVMEACPGAGSGAVGQOKELPKPKKTPPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60				
Db	1	MAVMEACPGAGSGAVGQOKELPKPKKTPPLGKKQSSVYKLEAVEKSPVFCGKWEILND	60				
Qy	61	VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSRIFIAAGSKQYQSSESLOIPNN	120				
Db	61	VITKGTAKGSEAGPAAISIIAQACENSQSFPTFSRIFIAAGSKQYQSSESLOIPNN	120				
Qy	121	VAHATEGMARVCKGKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE	180				
Db	121	VAHATEGMARVCKGKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE	180				
Qy	181	DESLPLGAPYVNTQFTKPLKEPGLQCFKQLEGRLPALPRSLHKLISPLQCLNHVW	240				
Db	181	DESLPLGAPYVNTQFTKPLKEPGLQCFKQLEGRLPALPRSLHKLISPLQCLNHVW	240				
Qy	241	KLHPDQGGPLPPTHPPYSLRPLPFPHPKPHLPESFLGKLACVDSQKPLPDHP	300				
Db	241	KLHPDQGGPLPPTHPPYSLRPLPFPHPKPHLPESFLGKLACVDSQKPLPDHP	300				
Qy	301	LSKLACVDSKPLPGLPHLEPSCLSRGAHEKFSVEEYLVAHQGVSSSQHSLTSLAKTW	360				
Db	301	LSKLACVDSKPLPGLPHLEPSCLSRGAHEKFSVEEYLVAHQGVSSSQHSLTSLAKTW	360				



Query Match	99.8%;	Score 5044;	DB 2;	Length 947;
Best Local Similarity	99.9%;	Pred. No. 0;		

Matches	946;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	MVMEACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND	60						
Db	1	MVMEACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND	60						
Qy	61	VITGTAKGSEAGPAAISIIAQAECENSQEFSTFSERIFIAGSKQYSQSESLDQIPNN	120						
Db	61	VITGTAKGSEAGPAAISIIAQAECENSQEFSTFSERIFIAGSKQYSQSESLDQIPNN	120						
Qy	121	VAHATEGMARVCWKGRSKARKRKXKSSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180						
Db	121	VAHATEGMARVCWKGRSKARKRKXKSSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180						
Qy	181	DESLGAPYVRNTPQFTKPLKEPGLQCFKQGEGLRPALEPSELHKLISPLQCLNHVW	240						
Db	181	DESLGAPYVRNTPQFTKPLKEPGLQCFKQGEGLRPALEPSELHKLISPLQCLNHVW	240						
Qy	241	KLHPDQGGPLPLTHPPFYSRLPFPHPLOPKPKHPLESPLGKACVDSQKPLDPDH	300						
Db	241	KLHPDQGGPLPLTHPPFYSRLPFPHPLOPKPKHPLESPLGKACVDSQKPLDPDH	300						
Qy	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360						
Db	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360						
Qy	361	AAGRSRSPSKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420						
Db	361	AAGRSRSPSKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420						
Qy	421	QTGFQCAVKKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGOL	480						
Db	421	QTGFQCAVKKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGOL	480						
Qy	481	VKEQGCLPDRALYLGOALEGLEVYLSRRILHGDVKADNVLLSDGSHAALCDFGHAVC	540						
Db	481	VKEQGCLPDRALYLGOALEGLEVYLSRRILHGDVKADNVLLSDGSHAALCDFGHAVC	540						
Qy	541	LQPDGLGKSLITGDIYPTGTEHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTOFF	600						
Db	541	LQPDGLGKSLITGDIYPTGTEHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTOFF	600						
Qy	601	RGPLCLKIASPPPPVREIPPSCAPLTAQAIQEGLRKEPIHRYSAALGSGKNRALQQVGG	660						
Db	601	RGPLCLKIASPPPPVREIPPSCAPLTAQAIQEGLRKEPIHRYSAALGSGKNRALQQVGG	660						
Qy	661	LKSPWGEYKEPRHPPPNOANYHOTLHAQPRELSRAPGPRPAEETTGAPKLPPLPPE	720						
Db	661	LKSPWGEYKEPRHPPPNOANYHOTLHAQPRELSRAPGPRPAEETTGAPKLPPLPPE	720						
Qy	721	PPPNKSPPLTLKSBSGMEWELPLSSLEPAPARNPSPERKATVPPELOOLELEFLN	780						
Db	721	PPPNKSPPLTLKSBSGMEWELPLSSLEPAPARNPSPERKATVPPELOOLELEFLN	780						
Qy	781	SLSQPFLBEQOILISCLSIDSLSDSSEKNPKSKASSRDTLSSGVHSSWSSQAEARSS	840						
Db	781	SLSQPFLBEQOILISCLSIDSLSDSSEKNPKSKASSRDTLSSGVHSSWSSQAEARSS	840						
Qy	841	SNMVLARGRPDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS	900						
Db	841	SNMVLARGRPDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS	900						
Qy	901	LVTKDQGPVRYDMEVPDSDGIDLAQCTLAPDGSFAMSWRVKHGQLENRP	947						
Db	901	LVTKDQGPVRYDMEVPDSDGIDLAQCTLAPDGSFAMSWRVKHGQLENRP	947						

RESULT 5  
US-09-032-475-2  
; Sequence 2, Application US/09032475  
; Patent No. 5854003  
; GENERAL INFORMATION:  
; APPLICANT: Roche, Mike

; APPLICANT: Wu, Lin  
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032,475  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/887,518  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: T97-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 947 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-032-475-2

Query Match 99.8%; Score 5044; DB 2; Length 947;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MVMEACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND	60
Db	1	MVMEACPGAPGSAVGQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND	60
Qy	61	VITGTAKGSEAGPAAISIIAQAECENSQEFSTFSERIFIAGSKQYSQSESLDQIPNN	120
Db	61	VITGTAKGSEAGPAAISIIAQAECENSQEFSTFSERIFIAGSKQYSQSESLDQIPNN	120
Qy	121	VAHATEGMARVCWKGRSKARKRKXKSSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180
Db	121	VAHATEGMARVCWKGRSKARKRKXKSSKSLAHAGVALAKPLRTPPEQESCTIPVOE	180
Qy	181	DESLGAPYVRNTPQFTKPLKEPGLQCFKQGEGLRPALEPSELHKLISPLQCLNHVW	240
Db	181	DESLGAPYVRNTPQFTKPLKEPGLQCFKQGEGLRPALEPSELHKLISPLQCLNHVW	240
Qy	241	KLHPDQGGPLPLTHPPFYSRLPFPHPLOPKPKHPLESPLGKACVDSQKPLDPDH	300
Db	241	KLHPDQGGPLPLTHPPFYSRLPFPHPLOPKPKHPLESPLGKACVDSQKPLDPDH	300
Qy	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360
Db	301	LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW	360
Qy	361	AAGRSRSPSKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420
Db	361	AAGRSRSPSKPTEDNEGVLITKLPVDYEVREEVHWATHQLRGRGSGFGEVHRMEDK	420
Qy	421	QTGFQCAVKKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGOL	480
Db	421	QTGFQCAVKKVRLEVFRAEELMACAGLTSPRIPLYGAVREGPWVNFMELEGGSLGOL	480



QY 481 VKQCCLPEDRALYYLGOALEGLEYLHSHRRILHGDVADNVLSSDGSAAALCDPFGHVC 540  
Db 481 VKQCCLPEDRALYYLGOALEGLEYLHSHRRILHGDVADNVLSSDGSAAALCDPFGHVC 540  
QY 541 LOPDGLGSLTGDYIPCTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTQFF 600  
Db 541 LOPDGLGSLTGDYIPCTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTQFF 600  
QY 601 RGPLCKIASPPPPVREIPPPSCAPLTAQAIQEGRLKEPIHRVSAELGKVNRAALQVGG 660  
Db 601 RGPLCKIASPPPPVREIPPPSCAPLTAQAIQEGRLKEPIHRVSAELGKVNRAALQVGG 660  
QY 661 LKSPWRGYSKPRHPPPNQANHYOTLHAQPRELSRAPGPRPAETTGRAPKLQPPPE 720  
Db 661 LKSPWRGYSKPRHPPPNQANHYOTLHAQPRELSRAPGPRPAETTGRAPKLQPPPE 720  
QY 721 PPERNKSPPLTSLKEESGMBEPLPLSSLEPAPARNPSSPERKATVPQEQLQLELFLN 780  
Db 721 PPERNKSPPLTSLKEESGMBEPLPLSSLEPAPARNPSSPERKATVPQEQLQLELFLN 780  
QY 781 SLSQPFSLKEEQILSLCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSQAEARSS 840  
Db 781 SLSQPFSLKEEQILSLCLSIDSLSDSEKNPKSKASQSSRDTLSSGVHSSQAEARSS 840  
QY 841 SNNVYLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISSQIPAAAFS 900  
Db 841 SNNVYLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREPHRVKVGDIATGISSQIPAAAFS 900  
QY 901 LVTKDQGVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHQLENRP 947  
Db 901 LVTKDQGVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKHQLENRP 947

RESULT 6  
US-09-513-999C-6660  
; Sequence 6660, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/513.999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 6660  
; LENGTH: 94  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-513-999C-6660

Query Match 9.5%; Score 479; DB 4; Length 94;  
Best Local Similarity 98.9%; Pred. No. 6.4e-29;  
Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MAVMEMCPGAGSAGVQKLPKPKETPLGKKQSSVYKLEAVKSPVFCGKWEILND 60  
Db 1 MAVMEMCPGAGSAGVQKLPKPKETPLGKKQSSVYKLEAVKSPVFCGKWEILND 60  
QY 61 VITKGTAKGSEAGAPAAISIIAQAECENSQBFSP 94  
Db 61 VITKGTAKGSEAGAPAAISIIAQAECENSQBFSP 94

RESULT 7  
US-09-522-775A-4  
; Sequence 4, Application US/09522775A

; Patent No. 6660906  
; GENERAL INFORMATION:  
; APPLICANT: Tsichlis, Philip N.  
; TITLE OF INVENTION: Inhibition of Tpl2 To Treat Inflammatory Diseases  
; FILE REFERENCE: OTT-3202  
; CURRENT APPLICATION NUMBER: US/09/522.775A  
; CURRENT FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Rat Tpl2 polypeptide sequence  
US-09-522-775A-4  
Query Match 7.1%; Score 360.5; DB 4; Length 467;  
Best Local Similarity 27.9%; Pred. No. 5.3e-19;  
Matches 107; Conservative 66; Mismatches 127; Indels 83; Gaps 13;  
QY 332 SVEEYLVHALQGSVSSOASHLSLAKTWAARGSRSPSKTEDNEGVLLEKLPVDY 391  
Db 80 TVEDLL--AFANHISNTKHFY-----GCRPQE-----SGILLNMVISPQNG 119  
QY 392 EYREE-----VHWATHQLRLG-----RGSFGEVHRMEDKQTFQCAVKKVLEVFRABEL 441  
Db 120 RYQIDSDVLLVPWKLTYRSIGSGFVPRGAFGVLAQDMKTKRMACKLIPVDQFKPSDV 179  
QY 442 MACAGLTSPIVPLYGAVREGPWNIWELLEGSLQOLVKEQGLPDRALYVQLALE 501  
Db 180 EIQACFRHENIAELYGAVLWMDTVHLEMEAGEGSLVLEKLESCGPMREFEIIVTWKHLK 239  
QY 502 GLEYLHSHRRILHGDVADNVLSSDGSAAALCDPFGHVCVCLQPDGLGSLTGDYIP---- 557  
Db 240 GLDFLHKKVKIHHDIKPSNIVFMS--TKAVLVDFGLSVQMTED-----VTLPKDLR 288  
QY 558 GTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTQFF---RGPLCKIA-SEPP 613  
Db 289 GTEIYMSPEVILCRGHSTKADIYSLGATLIHMOTGTPPWVKRYPRSAVPSVLYIHKQAP 348  
QY 614 PVREIPSPALTAQAIQEGRLKEPIHRVSAELGKVNRAALQVGGGLKSPWGEYKEPR 673  
Db 349 PLEDIAGDCSPGMRLEIAALERNPNRPKAADL-----LK----- 384  
QY 674 HPPNQANYHOTLHAQPRELSR 696  
Db 385 -----HEALN-PPREDQPR 397

RESULT 8  
US-09-522-775A-2  
; Sequence 2, Application US/09522775A  
; Patent No. 6660906  
; GENERAL INFORMATION:  
; APPLICANT: Tsichlis, Philip N.  
; TITLE OF INVENTION: Inhibition of Tpl2 To Treat Inflammatory Diseases  
; FILE REFERENCE: OTT-3202  
; CURRENT APPLICATION NUMBER: US/09/522.775A  
; CURRENT FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human Tpl2 polypeptide sequence  
US-09-522-775A-2

Query Match 7.1%; Score 360; DB 4; Length 467;  
Best Local Similarity 27.2%; Pred. No. 5.8e-19;  
Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;







Db 203 ---SPVFTQSRPPSSNIHRPKPSRPVP-----GSTSKLGDAKSMMLDLGSA 248  
Qy 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEYLVAHQSVSS-- 347  
Db 249 SRCDDSFGGGNGNAVIPSDETPTVEDKRLDVTNELNSSIEDLLEASMPSSDITVT 308  
Qy 348 -----SOAHSITSLAKTWAARGSR--EPSP 372  
Db 309 FKSEVAVLSPEKAENDDTYKDDVNNHKNQCKEKEEAEALAIAMWSASQDALPIV 368  
Qy 373 KTEDEGVLLEKLP-----VDYREBEVHWATHQLRGRSGFGEVHRMEDKOTGF 424  
Db 369 QVENGEDIIIIQDDTPETLPCHTKAKQPYREDAEWLKGQ-QIGLGFSSCYQAQDVGTGT 427  
Qy 425 QCAVKV-----RLEVFA--EELMACAGLTSPRIVLYGAVREGPWNIFMELLE 473  
Db 428 LMAVKQVTVYRNTSSEQEVEEALREIRMGHLNHPNIIRMLGATCEKSNYNLFIEWMA 487  
Qy 474 GGSGLQVLKQCGCLPEDRALYILGQALEGLYHLSRRILHGDVADNVLLSSDGSAAALC 533  
Db 488 GGSVAHLLSKYGAFKESVINYTEQLRLGLSYLHENQIHRDVKGANLLIDSTGQRLIA 547  
Qy 534 DFHGAVALQPDGLGKSLTGDYIPCTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593  
Db 548 DFGAARLASKGTGAGEFQGLL-GTIAFMAPEVLRGQYGRSCDVMSVGCATTEMACAK 606  
Qy 594 HPWTQFFRG---PLCLKIASE-----PPVREIPPPSCAPLTAQ--AIQGLRKEP 638  
Db 607 PPWNAEKHSNHLALIFKIASATTAPSIHSLSPGLRDVAVRCLQLQPDPPSRELLKHP 666  
Qy 639 IHRVS 643  
Db 667 VFRTT 671

RESULT 14

US-08-461-146C-2  
; Sequence 2, Application US/08461145C  
; Patent No. 5981265  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, GARY L.  
; TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lahive and Cockfield  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,146C  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/354,516  
; FILING DATE: 21-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/049,254  
; FILING DATE: 15-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/323,460  
; FILING DATE: 14-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/11690  
; FILING DATE: 14-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/04178

RESULT 15

US-08-461-145C-2  
; Sequence 2, Application US/08461145C  
; Patent No. 6074861

; FILING DATE: 15-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KARA, Catherine J.  
; REGISTRATION NUMBER: P41,106  
; REFERENCE/DOCKET NUMBER: CPI-004CN3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 672 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-461-146C-2  
  
Query Match 7.0%; Score 355.5; DB 2; Length 672;  
Best Local Similarity 23.6%; Pred. No. 2.1e-18;  
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;  
  
Qy 77 AISIIAQAECENSQEFSPFTSERIFIAGSKQYSQSESLDQIPNNVAHATEGKMARVCWK 136  
Db 58 SLQAVATPSCLENSLSLHTVHREKTGKLSATRLSASSEDISDRLAGVSVGLPSSTTTTEQ 117  
Qy 137 KR---RSKARKRKKSLSLAHAGVALAKPLPRTPEQESCTIPVQEDESPGLGAPYVNT 193  
Db 118 PKPAVQTKGRPHSQCLNSSPLSHA--QLMFPAPAPCSPSPV---DIS-----KHR 165  
Qy 194 POFTKPLKEGQLGQCFKQLEGRLPALPRSELHKLISPLQ--CLNHWKLLHHDDGGPL 251  
Db 166 POAFVPCIPS-----ASPTQ--RKFSLQFORNCSEH-----RDSQL 202  
Qy 252 PLTPHPFPYRLPHFPFPHLPQPMKPHLSFLKACVDSQKPLPDPHLSKLA----- 305  
Db 203 ---SPVFTQSRPPSSNIHRPKPSRPVP-----GSTSKLGDAKSMMLDLGSA 248  
Qy 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEYLVAHQSVSS-- 347  
Db 249 SRCDDSFGGGNGNAVIPSDETPTVEDKRLDVTNELNSSIEDLLEASMPSSDITVT 308  
Qy 348 -----SOAHSITSLAKTWAARGSR--EPSP 372  
Db 309 FKSEVAVLSPEKAENDDTYKDDVNNHKNQCKEKEEAEALAIAMWSASQDALPIV 368  
Qy 373 KTEDEGVLLEKLP-----VDYREBEVHWATHQLRGRSGFGEVHRMEDKOTGF 424  
Db 369 QVENGEDIIIIQDDTPETLPCHTKAKQPYREDAEWLKGQ-QIGLGFSSCYQAQDVGTGT 427  
Qy 425 QCAVKV-----RLEVFA--EELMACAGLTSPRIVLYGAVREGPWNIFMELLE 473  
Db 428 LMAVKQVTVYRNTSSEQEVEEALREIRMGHLNHPNIIRMLGATCEKSNYNLFIEWMA 487  
Qy 474 GGSGLQVLKQCGCLPEDRALYILGQALEGLYHLSRRILHGDVADNVLLSSDGSAAALC 533  
Db 488 GGSVAHLLSKYGAFKESVINYTEQLRLGLSYLHENQIHRDVKGANLLIDSTGQRLIA 547  
Qy 534 DFHGAVALQPDGLGKSLTGDYIPCTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593  
Db 548 DFGAARLASKGTGAGEFQGLL-GTIAFMAPEVLRGQYGRSCDVMSVGCATTEMACAK 606  
Qy 594 HPWTQFFRG---PLCLKIASE-----PPVREIPPPSCAPLTAQ--AIQGLRKEP 638  
Db 607 PPWNAEKHSNHLALIFKIASATTAPSIHSLSPGLRDVAVRCLQLQPDPPSRELLKHP 666  
Qy 639 IHRVS 643  
Db 667 VFRTT 671



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 12:02:31 ; Search time 119.819 Seconds  
(without alignments)  
2846.907 Million cell updates/sec

Title: US-09-155-676B-7  
Perfect score: 5052  
Sequence: 1 MAVMEMACPGAPGSAVQKQK.....PDGSFAWSRVKIGQLNRP 947

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5052	100.0	947	9 US-09-871-889-1	Sequence 1, Appli
2	5052	100.0	947	10 US-09-981-397A-18	Sequence 18, Appl
3	5052	100.0	947	15 US-10-394-322A-44	Sequence 44, Appl
4	5040	99.8	947	13 US-10-087-192-888	Sequence 888, App
5	2481.5	49.1	548	13 US-10-087-192-888	Sequence 885, App
6	369	7.3	891	15 US-10-425-114-57758	Sequence 57758, A
7	367	7.3	250	10 US-09-976-782-31	Sequence 31, Appl
8	367	7.3	250	15 US-10-042-865-102	Sequence 102, App
9	367	7.3	251	15 US-10-052-648A-41	Sequence 41, Appl
10	367	7.3	252	15 US-10-074-978A-160	Sequence 160, App
11	367	7.3	254	15 US-10-074-978A-150	Sequence 150, App
12	367	7.3	255	15 US-10-087-684-98	Sequence 98, Appl

13	367	7.3	255	15	US-10-218-779-98	Sequence 98, Appl
14	367	7.3	256	11	US-09-964-956-72	Sequence 72, Appl
15	367	7.3	256	15	US-10-055-569A-99	Sequence 99, Appl
16	367	7.3	256	15	US-10-042-865-83	Sequence 83, Appl
17	367	7.3	256	15	US-10-072-012-799	Sequence 799, App
18	367	7.3	256	15	US-10-072-012-856	Sequence 856, App
19	367	7.3	256	15	US-10-072-012-876	Sequence 876, App
20	367	7.3	256	15	US-10-072-012-882	Sequence 882, App
21	367	7.3	256	15	US-10-029-020-183	Sequence 183, App
22	360	7.1	467	15	US-10-287-226-24	Sequence 24, Appl
23	360	7.1	467	15	US-10-287-226-36	Sequence 36, Appl
24	360	7.1	468	15	US-10-287-226-8	Sequence 8, Appli
25	360	7.1	471	15	US-10-287-226-18	Sequence 18, Appl
26	360	7.1	474	15	US-10-287-226-4	Sequence 4, Appli
27	360	7.1	474	15	US-10-287-226-26	Sequence 26, Appl
28	360	7.1	475	15	US-10-287-226-16	Sequence 16, Appl
29	360	7.1	475	15	US-10-287-226-32	Sequence 32, Appl
30	360	7.1	478	15	US-10-287-226-10	Sequence 10, Appl
31	360	7.1	486	15	US-10-287-226-20	Sequence 20, Appl
32	360	7.1	486	15	US-10-287-226-22	Sequence 22, Appl
33	359	7.1	351	16	US-10-664-421-155	Sequence 155, App
34	359	7.1	467	9	US-09-374-579-4	Sequence 4, Appli
35	359	7.1	467	15	US-10-374-414-4	Sequence 4, Appli
36	359	7.1	467	15	US-10-262-511-228	Sequence 228, App
37	359	7.1	467	15	US-10-287-226-2	Sequence 2, Appli
38	358.5	7.1	1302	13	US-10-000-864-2	Sequence 2, Appli
39	358.5	7.1	1511	15	US-10-433-794-6	Sequence 6, Appli
40	358.5	7.1	1511	16	US-10-618-941-109	Sequence 109, App
41	358	7.1	256	10	US-09-863-776-40	Sequence 40, Appl
42	357	7.1	1512	15	US-10-210-130-64	Sequence 64, Appl
43	356.5	7.1	423	15	US-10-287-226-30	Sequence 30, Appl
44	356.5	7.1	424	15	US-10-287-226-34	Sequence 34, Appl
45	356.5	7.1	425	15	US-10-287-226-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-871-889-1  
; Sequence 1, Application US/09871889  
; Patent No. US20020042499A1  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Warner C.  
; APPLICANT: Lin, Xin  
; APPLICANT: Gelezuinas, Romas  
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED  
; FILE REFERENCE: 30448.61USUI  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: US/09/871.889  
; PRIOR FILING DATE: 1999-02-25  
; PRIOR APPLICATION NUMBER: 60/076,299  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)  
US-09-871-889-1

Query Match	100.0%;	Score 5052;	DB 9;	Length 947;
Best Local Similarity	100.0%;	Pred. No. 1.8e-286;		
Matches 947;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAVMEACPGAPGSAVQKQELPKPKETPLGKKQSSVVKLEAVEKSPVFCGKWEI	ND	60
Db	1	MAVMEACPGAPGSAVQKQELPKPKETPLGKKQSSVVKLEAVEKSPVFCGKWEI	ND	60
Qy	61	VITKGTAKESGAPPAISIIAQAECSNQESPTFSERIFTAGSKQYSQESLDDQIPNN		120

```
Db 61 VITKTAKGSEAGPAAISIIAQAECENSQEFSTFSEIRIFTAGSKQYSQSESLOIINN 120
QY 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPQEOSTTIPVOE 180
Db 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPQEOSTTIPVOE 180
QY 181 DESPLGAPVYRNTPOFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVYRNTPOFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHPDQGGPLPLTHPPFPYSLRPLPFPPLPQWPKHPLESPLGKLACVDQSKPLPDH 300
Db 241 KLHPDQGGPLPLTHPPFPYSLRPLPFPPLPQWPKHPLESPLGKLACVDQSKPLPDH 300
QY 301 LSKLACVDSPKPLGPHLPSPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSITSLAKTW 360
Db 301 LSKLACVDSPKPLGPHLPSPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSITSLAKTW 360
QY 361 AAGRSRSPSKPTEDNEGVLITLTKLPVDYREVEVHWATHQLRLGRGSFGEVHRMEDK 420
Db 361 AAGRSRSPSKPTEDNEGVLITLTKLPVDYREVEVHWATHQLRLGRGSFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYILGOALEGLYLHRSRRI LHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYILGOALEGLYLHRSRRI LHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGKVNRAALQOYVG 660
Db 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGKVNRAALQOYVG 660
QY 781 SLSPFSLSEOEQILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSQAARSS 840
Db 781 SLSPFSLSEOEQILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSQAARSS 840
QY 841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900
Db 841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900
QY 901 LVTQKQPVRYDMEVPDSDGIDLOCTLAPDGSFAWSRWVKHGLENNR 947
Db 901 LVTQKQPVRYDMEVPDSDGIDLOCTLAPDGSFAWSRWVKHGLENNR 947
```

RESULT 2

```
US-09-981-397A-18
; Sequence 18, Application US/09981397A
; Publication No. US20030082519A1
; GENERAL INFORMATION:
; APPLICANT: Axxima Pharmaceuticals AG
; APPLICANT: Schubart, Daniel
; APPLICANT: Habenberger, Peter
; APPLICANT: Stein-Gerlach, Matthias
; APPLICANT: Bevec, Dorian
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their
; TITLE OF INVENTION: Inhibition
; FILE REFERENCE: AXM-004.1 US
; CURRENT APPLICATION NUMBER: US/09/981,397A
```

```
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/240,750
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-397A-18
```

```
Query Match 100.0%; Score 5052; DB 10; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.8e-286;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQOKELPKPKKKTTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQOKELPKPKKKTTPPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKTAKGSEAGPAAISIIAQAECENSQEFSTFSEIRIFTAGSKQYSQSESLOIINN 120
Db 61 VITKTAKGSEAGPAAISIIAQAECENSQEFSTFSEIRIFTAGSKQYSQSESLOIINN 120
QY 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPQEOSTTIPVOE 180
Db 121 VAHATGKMARVCWKGRSKARKKSKSLAHAGVALAKPLRTPQEOSTTIPVOE 180
QY 181 DESPLGAPVYRNTPOFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPVYRNTPOFTKPLKEPGLGQLCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHPDQGGPLPLTHPPFPYSLRPLPFPPLPQWPKHPLESPLGKLACVDQSKPLPDH 300
Db 241 KLHPDQGGPLPLTHPPFPYSLRPLPFPPLPQWPKHPLESPLGKLACVDQSKPLPDH 300
QY 301 LSKLACVDSPKPLGPHLPSPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSITSLAKTW 360
Db 301 LSKLACVDSPKPLGPHLPSPCLSRGAHEKFSVEEYLVAHALQGSVSSQAHSITSLAKTW 360
QY 361 AAGRSRSPSKPTEDNEGVLITLTKLPVDYREVEVHWATHQLRLGRGSFGEVHRMEDK 420
Db 361 AAGRSRSPSKPTEDNEGVLITLTKLPVDYREVEVHWATHQLRLGRGSFGEVHRMEDK 420
QY 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIRVPLYGAVREGPWVNI FMELEGGSLGOL 480
QY 481 VKEQGCLPEDRALYILGOALEGLYLHRSRRI LHGDVKADNVLLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGCLPEDRALYILGOALEGLYLHRSRRI LHGDVKADNVLLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTOFF 600
QY 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGKVNRAALQOYVG 660
Db 601 RGPLCLKIASEPPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAEELGKVNRAALQOYVG 660
QY 661 LKSPWRGEYKEPRHPPPNQAQYHOTLHAQPRELSPRAPGPRPAEETTGRAKLOPPLPPE 720
Db 661 LKSPWRGEYKEPRHPPPNQAQYHOTLHAQPRELSPRAPGPRPAEETTGRAKLOPPLPPE 720
QY 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPPELOLQLEIEFLN 780
Db 721 PPEPNKSPPLTLSKEESGMWEPLPLSSLEPAPARNPSSPERKATVPPELOLQLEIEFLN 780
QY 781 SLSPFSLSEOEQILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSQAARSS 840
Db 781 SLSPFSLSEOEQILSCLSIDSLSDSEKKNPSKASQSSRDTLSSGVHSSWSQAARSS 840
QY 841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900
Db 841 SWNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSIQIPAAAFS 900
```



```
Db 841 SNNVLAGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGTGISSQIPAAAFS 900
Qy 901 LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
Db 901 LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947

RESULT 3
US-10-394-322A-44
; Sequence 44, Application US/10394322A
; Publication No. US20030232391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-44

Query Match 100.0%; Score 5052; DB 15; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.8e-286;
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGATKEGSEAGPAALIIIAQACENSQSFSTFSEIRIFIAQSKQYQSESIDQIPNN 120
Db 61 VITKGATKEGSEAGPAALIIIAQACENSQSFSTFSEIRIFIAQSKQYQSESIDQIPNN 120

Qy 121 VAHATEGMARVCKWKGKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180
Db 121 VAHATEGMARVCKWKGKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180

Qy 181 DESPLGAPYVRNTPOFTKPLKEPGLQCFKQEGELRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLQCFKQEGELRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHPQDGGPLPLPHTPPYSLRPLPFPFHPQPKWPHLESFLGKLACVDSQKPLDPH 300
Db 241 KLHPQDGGPLPLPHTPPYSLRPLPFPFHPQPKWPHLESFLGKLACVDSQKPLDPH 300

Qy 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKESVVEYLVAHQSGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKESVVEYLVAHQSGSVSSQAHSLSLAKTW 360

Qy 361 AARGSRREPSPKTDNEGVLLTEKLKPDVDEYEEVHWATHQLRGLGRSGFGEVHRMEDK 420
Db 361 AARGSRREPSPKTDNEGVLLTEKLKPDVDEYEEVHWATHQLRGLGRSGFGEVHRMEDK 420

Qy 421 QTGFQCAVKVRLEVFRAELMACAGLTSRPRIVLYGAVREGPWVNIIFMELLEGSLGQL 480
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSRPRIVLYGAVREGPWVNIIFMELLEGSLGQL 480

Qy 481 VKEQCLPEDRALYLGQALEGLYLSRRRIHGDVKADNVLLSDGSHALCDFGHAVC 540
Db 481 VKEQCLPEDRALYLGQALEGLYLSRRRIHGDVKADNVLLSDGSHALCDFGHAVC 540

Qy 541 LQPDGLGSLTLDGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
Db 541 LQPDGLGSLTLDGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600

Qy 601 RGPLCLTIASRPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKVNRLAQVGG 660
```

```
Db 601 RGPLCLTIASRPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAELGKVNRLAQVGG 660
Qy 661 LKSPWRGEGYKEPRHPPPNQANYHTLHAQRELSPRAPGPRPAEETTRAPKIQPLPPE 720
Db 661 LKSPWRGEGYKEPRHPPPNQANYHTLHAQRELSPRAPGPRPAEETTRAPKIQPLPPE 720

Qy 721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQOLEIEFLN 780
Db 721 PPEPNKSPPLTSLKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQOLEIEFLN 780

Qy 781 SLSQPFSLERQEQILSCLSIDSLSDSEKXNPKASQSSRDITLSSGVHSSQAEARSS 840
Db 781 SLSQPFSLERQEQILSCLSIDSLSDSEKXNPKASQSSRDITLSSGVHSSQAEARSS 840

Qy 841 SNNVLAGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGTGISSQIPAAAFS 900
Db 841 SNNVLAGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAIGTGISSQIPAAAFS 900

Qy 901 LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
Db 901 LVTKDQPVRYDMEVPSGIDLQCTLAPDGSFAMSWRVKHGQLENRP 947
```

## RESULT 4

```
US-10-087-192-888
; Sequence 888, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 888
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-888
```

```
Query Match 99.8%; Score 5040; DB 13; Length 947;
Best Local Similarity 99.8%; Pred. No. 9.2e-286;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAVMEMACPGAGSAGVQKQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGATKEGSEAGPAALIIIAQACENSQSFSTFSEIRIFIAQSKQYQSESIDQIPNN 120
Db 61 VITKGATKEGSEAGPAALIIIAQACENSQSFSTFSEIRIFIAQSKQYQSESIDQIPNN 120

Qy 121 VAHATEGMARVCKWKGKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180
Db 121 VAHATEGMARVCKWKGKRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCTIPVOE 180

Qy 181 DESPLGAPYVRNTPOFTKPLKEPGLQCFKQEGELRPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLQCFKQEGELRPALPRSELHKLISPLQCLNHVW 240

Qy 241 KLHPQDGGPLPLPHTPPYSLRPLPFPFHPQPKWPHLESFLGKLACVDSQKPLDPH 300
Db 241 KLHPQDGGPLPLPHTPPYSLRPLPFPFHPQPKWPHLESFLGKLACVDSQKPLDPH 300

Qy 301 LSKLACVDSQKPLDPHLEPSCLSRGAHEKESVVEYLVAHQSGSVSSQAHSLSLAKTW 360
```

Db 301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSGQAHSLTSLAKTW 360  
Qy 361 AARGSSRSPKTEDENEGVLTEKLPVDYEVREEVHATHQLRLGRGSFGEVHRMEDK 420  
Db 361 AARGSSRSPKTEDENEGVLTEKLPVDYEVREEVHATHQLRLGRGSFGEVHRMEDK 420  
Qy 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIIVPLYGAVREGPVWNI FMELEGGSLGQL 480  
Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIIVPLYGAVREGPVWNI FMELEGGSLGQL 480  
Qy 481 VKEQGLCPEDRALYYLGOALEGLEYLHRSRILHGDVKNVLLSDGSHAAALCDFGHAYC 540  
Db 481 VKEQGLCPEDRALYYLGOALEGLEYLHRSRILHGDVKNVLLSDGSHAAALCDFGHAYC 540  
Qy 541 LQPDGGLKSLTGTGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTPFF 600  
Db 541 LQPDGGLKSLTGTGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTPFF 600  
Qy 601 RGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRYSAABELGKGNRALQOVGG 660  
Db 601 RGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRYSAABELGKGNRALQOVGG 660  
Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPPRAGPRPAEETTGRAKLPPLPPE 720  
Db 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPPRAGPRPAEETTGRAKLPPLPPE 720  
Qy 721 PPEPNKSPPLTUSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLELEFLN 780  
Db 721 PPEPNKSPPLTUSKEESGMWEPPLSSLEPAPARNPSSPERKATVPEQELQLELEFLN 780  
Qy 781 SLSPQFLEEQILSCLSIDSLSDSEKPNKSPKASQSSRDTLSSGVHSSWSQAEARSS 840  
Db 781 SLSPQFLEEQILSCLSIDSLSDSEKPNKSPKASQSSRDTLSSGVHSSWSQAEARSS 840  
Qy 841 SNWMLARGRPDTSPYFNGVKVQSLNGEHLHIREFHVRKVGDIATGISSQIPAAAFS 900  
Db 841 SNWMLARGRPDTSPYFNGVKVQSLNGEHLHIREFHVRKVGDIATGISSQIPAAAFS 900  
Qy 901 LVTKGQPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENRP 947  
Db 901 LVTKGQPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENRP 947

## RESULT 5

US-10-087-192-885  
; Sequence 885, Application US/10087192  
; Publication No. US20020182386A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 885  
; LENGTH: 548  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-087-192-885

Query Match 49.1%; Score 2481.5; DB 13; Length 548;  
Best Local Similarity 86.1%; Pred. No. 1.2e-136;  
Matches 466; Conservative 32; Mismatches 36; Indels 7; Gaps 2;

Qy 407 GRGSFGEVHRMEDKGTGFQCAVKVRLEVFRAELMACAGLTSPRIIVPLYGAVREGPVWNI 466

Db 15 GQSSFGEVHRMEDKGTGFQCAVKVRLEVFRAELMACAGLTSPRIIVPLYGAVREGPVWNI 74  
Qy 467 IFMELLEGSLGQLVKBOGCLPEDRALYYLGOALEGLEYLHRSRILHGDVKNVLLSSD 526  
Db 75 IFMELLEGSLGQLVKBOGCLPEDRALYYLGOALEGLEYLHRSRILHGDVKNVLLSSD 134  
Qy 527 GSHAALCDFGHAYCLOPDGLGKSLTGTGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCM 586  
Db 135 GSHAALCDFGHAYCLOPDGLGKSLTGTGYIPGTETHMAPEVVLGRSCDAKVDVWSSCCM 194  
Qy 587 LHMINGCHPWTPQFFRGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRYSAAE 646  
Db 195 LHMINGCHPWTPQFFRGPLCLKIASPPVREIIPSCAPLTAQAIQEGLRKEPIHRYSAAE 254  
Qy 647 LGKGNRALQOVGGUKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPPRAGPRPAEET 706  
Db 255 LRRKVGKALQEVGGLKSPWKGEYKEPRPPQDQATCHOTLPTPRE-NP-----PAKAN 307  
Qy 707 TGRAPKLOPPLPPEPPEPNKSPPLTUSKEESGMWEPPLSSLEPAPARNPSSPERKATV 766  
Db 308 TDGAPEPQPLPPEPPEPNKSPALMLSKESGTMWEPPLSSLDLPATAKGFPPDRATLP 367  
Qy 767 EQELQLELEFLNLSLSQFFSLEEQEILSCLSIDSLSDSEKPNKSPKASQSSRDTLSS 826  
Db 368 EQELQLELEFLNLSLSQFFSLEEQEILSCLSIDSLSDSEKPNKSPKASQSSRDTLSS 427  
Qy 827 GVHSSWSQAEARSSWNWMLARGRPDTSPYFNGVKVQSLNGEHLHIREFHVRKVGDI 886  
Db 428 GVHSSWSQAEARSSWNWMLARGRPDTSPYFNGVKVQSLNGEHLHIREFHVRKVGDI 487  
Qy 887 ATGSSQIPAAAFSLVTKGQPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENR 946  
Db 488 ATGSSQIPAAAFSLVTKGQPVRYDMEVPDSDGIDLOCTLAPDGSFAMSWRVKVGQLENR 547  
Qy 947 P 947  
Db 548 P 548

## RESULT 6

US-10-425-114-57758  
; Sequence 57758, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57758  
; LENGTH: 891  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOY034B02\_FLI.fli  
US-10-425-114-57758

Query Match 7.3%; Score 369; DB 15; Length 891;  
Best Local Similarity 23.4%; Pred. No. 4.5e-13;  
Matches 224; Conservative 107; Mismatches 353; Indels 272; Gaps 43;  
Qy 88 NSQSFSPFSEIRIFIA-----GSKQYSQ---SESIDQIPNNVAHATEGKMARVCW 134  
Db 21 NKESFFDTLHRKLRKISSKGVKHSKGSRRHRCNTTISEKGDHSPCGSRSPSPKVAR-CQ 79

QY 135 KGRSKAR-----KRRKKSSKSLAH-----AGVALAKPLP 166  
Db 80 SFIDRPHAPLPLGLHPSSVGRVDSEISISKSRLEKVRPLSLTLTPTGRCRCRNP 139  
QY 167 RTPEESCTIPQVED-----ESPL-----GAP---VVRNTPQFTK 198  
Db 140 ADLGDWTVASVFDSCSADSDPADSHNRSLAIDCETGTTAAGSPSSLMKQPPPAVS 199  
QY 199 PL-----KEPGLGOLCFKQGLGELRPLPRSELHKLISPLQCLNHWKMLHHHPQDGGPLPL 253  
Db 200 QLNSTGVKKPG-----NILSNHMSSTSPKRR-----PLR--NHVPNLQVPHGA--- 241  
QY 254 PTHFPYRSLRPHPPFFHPLQW-----KHPLESFLGKLACV----- 290  
Db 242 -FYSAPDSSLSRP--SRSLPRAFGTDQVLNSAFLAGKPYFEINFGSGHCSPGSHNSGH 299  
QY 291 -----DSQKPL---PDPHLSKLACVDSPK-PLPGPHLEPSCLSRGA-----HEKESVEEYL 337  
Db 300 NSMGDMGGLLWQPSRGSPSPVSPRMTSPGP---SSRIQSGAVTPIHPKAG----- 351  
QY 338 VHALQGSVSSSOAH-----SLTSLAKTAAAGRSRSPSPKTEDNEGVLLTEK 385  
Db 352 -----GTPTESQTHRLPLPLSVSNSSLSFHSNSAATSPSPRSPARADNPSG----- 400  
QY 386 LKPDVDEYREEVHATHOLRLGRSGFGEVHRMEDKQTFQCAVKKVRL-----E 434  
Db 401 -----SRWKKGKL--LGSGSGFHVILGFNSRGEMCAVKEVTLFSDDPKSMESAK 448  
QY 435 VFRAELMACAGLTSPRVPLYGAVRGPWNIFWELLEGSLGOLVKEQCLEDRALY 494  
Db 449 QF--MOEIHLSRLQHPNIVQYGETVDNKLIYIYLVYVSGSIHKLREYQFGELVIRS 507  
QY 495 YLQALGLELYLHSHRRILHGVKADNVLSSDGSGLHAALCDFGHAVCLQPDGLGKSLTGD 554  
Db 508 YTOQLSLGLYLHAKNTLHRDIKANILVDPTG--RVKLADFGMA-----KHTGQ 556  
QY 555 YIP-----GTEHMAPEVVL--GRSCDAKVDVWSSCMMHMLNGCHPWTQFFRGPLCLKIA 609  
Db 557 SCPLSFKGTYPYMAPEVIKNSGNCNLAVDIWSLGCIVLEMTATKPPWFQYEGVAAMFKIG 616  
QY 610 SEPPVREIPSCAPLTAQA---TOEGLRKEPIHRVSAEL-----GKYNR----- 653  
Db 617 NS-----KELTIPDHLNNEGDFVRKCLQRNPHRPSASELDDHPFVKNAAPLERPIAP 672  
QY 654 -ALQOVGGL---KSPWRGEYKEPRHPPNQAHTOT--LHAQPRELSRAPAGPRPAET 706  
Db 673 EALDPVSGITGAKALAIQGRNLSLSDSLSRSLVHSSRFLKTNPHESEIHP----- 724  
QY 707 TGRAPKLQPLPP--EPPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSSPERKAT 764  
Db 725 ---RNISCPVSPIGSPLLRSRP-----QHRNGKMSPSPISS--PRTASGASTP----- 768  
QY 765 VPBQLQOLELEFLNLSQFSLSEBQILUSCLSIDSLSDSEKKNPSKASOSSRDTL 824  
Db 769 -----LAGGSAIPFGNHSKQSIYFQEGFGSIIPKSSNGVMYMGHSHSDSNVDI 816  
QY 825 SSGVHSSSOAEARSSNNMVLAR---GRPTDTPSYNGVKVQI--QSLNGEHLHI 875  
Db 817 FRGMQSHSIPELVSSNDVLVQFAHPHAEPIYDFQSVLADRVGRQLLGEHVKI 872

RESULT 7  
US-09-76-782-31  
; Sequence 31, Application US/09976782  
; Publication No. US20030190715A1  
; GENERAL INFORMATION:  
; APPLICANT: Grosbe et al  
; TITLE OF INVENTION: No. US20030190715A1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-157  
; CURRENT APPLICATION NUMBER: US/09/976,782  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240,113  
; PRIOR FILING DATE: 2000-10-12

; PRIOR APPLICATION NUMBER: 60/240,662  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/240,732  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/240,625  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/240,703  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/241,190  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/240,637  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/240,669  
; PRIOR FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 60/262,455  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/240,648  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 31  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Consensus  
; OTHER INFORMATION: sequence  
US-09-976-782-31

Query Match 7.3%; Score 367; DB 10; Length 250;  
Best Local Similarity 37.9%; Pred. No. 1.3e-13;  
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRSGFGEVHRMEDKQTFQCAVKKVRLVFRAE-----ELMACAGLTSPRIVPLVGA 458  
Db 7 LGKAGFKVYLARDKTKGLVAKIRKEKLLKKGRRLIRKILKLDHPNIVKLTVDV 66  
QY 459 VREGFPWNIFWELLEGSLGOLVKEQCLEDRALYYLQALGLELYLHSHRRILHGDVKA 518  
Db 67 FEDDDKLYLVWEYCEGGDLFDLLKKRGLSEDEARFYARQILSALEYLHSGQIIHRDLKP 126  
QY 519 DNVLLSSDGSGLHAALCDFGHAVCLQPDGLGKSLTGDYIPGTETHMAPEVLGRSCDAKVD 578  
Db 127 ENILLSDSG--HVKLADFLGAKQLDSSG---TLIT--TFVGTPEYMAPEVLHGKYGKAVD 180  
QY 579 VWSSCMMHMLNGCHPWTQFFRG-----PLCLKIASPPPPVREIPP---SCAPLTAQAI 630  
Db 181 IWSLGVILYELLTGKPP-----FPGDDQLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233  
QY 631 QEGLRKEPIHRVSAE 646  
Db 234 KLLVKDPEKRLTAE 249

RESULT 8  
US-10-042-865-102  
; Sequence 102, Application US/10042865  
; Publication No. US20040029216A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Caeman, Stacie J  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Zhong, Mei  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Burgess, Catherine E  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Vernet, Corine A.M  
; APPLICANT: Taylor, Sarah  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Miller, Charles E



[illegible]

```
; PRIOR APPLICATION NUMBER: 60/312,284
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/268,496
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 60/276,703
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/330,293
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/322,127
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/280,899
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/310,797
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/268,646
; PRIOR FILING DATE: 2001-02-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 150
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: s_TKc,
; OTHER INFORMATION: Serine/Threonine protein kinases
US-10-074-978A-150

Query Match          7.3%; Score 367; DB 15; Length 254;
Best Local Similarity 37.9%; Pred. No. 1.3e-13;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRSGFGEVHRMEDKQTQCQAVKVRLEVFRAE-----ELMACAGLTSPRIVLYGA 458
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 7 LGKGAFGKYVLARDKTKGLVAIKVKIKKKRRLRIKILKLDHPNIVKLYDV 66

QY 459 VREGPWNVFMELLEGSLGQLVKEQGLPEDRALYLLQALLEGLEYLHRSRILHGDVKA 518
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 67 FEDDDKLYLVMEYCEGGDFDLKKRGRLSEDEARFYARQILSALEYLHSGQIIHRDLKP 126

QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVVLGRSCDAKD 578
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 127 ENILLDSGD-HVKLADFGAKQLDSGG---TLT--TFVGTPEYMAPEVLLKGYGKAVD 180

QY 579 VNSSCCMLHMLNGCHPWTFQFRG-----PLCLKIASPEPPVREIPP---SCAPLTAQAI 630
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 IWSGLVILVELLTGKPP-----FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233

QY 631 QEGLRKEPIHRVSAE 646
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 234 KLLVKDPEKRLTAEE 249

RESULT 12
US-10-087-684-98
; Sequence 98, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catharine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
```

```
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 98
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain
US-10-087-684-98

Query Match          7.3%; Score 367; DB 15; Length 255;
Best Local Similarity 37.9%; Pred. No. 1.4e-13;
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

QY 406 LGRSGFGEVHRMEDKQTQCQAVKVRLEVFRAE-----ELMACAGLTSPRIVLYGA 458
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 6 LGKGAFGKYVLARDKTKGLVAIKVKIKKKRRLRIKILKLDHPNIVKLYDV 65

QY 459 VREGPWNVFMELLEGSLGQLVKEQGLPEDRALYLLQALLEGLEYLHRSRILHGDVKA 518
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 66 FEDDDKLYLVMEYCEGGDFDLKKRGRLSEDEARFYARQILSALEYLHSGQIIHRDLKP 125

QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVVLGRSCDAKD 578
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 126 ENILLDSGD-HVKLADFGAKQLDSGG---TLT--TFVGTPEYMAPEVLLKGYGKAVD 179

QY 579 VNSSCCMLHMLNGCHPWTFQFRG-----PLCLKIASPEPPVREIPP---SCAPLTAQAI 630
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 180 IWSGLVILVELLTGKPP-----FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 232

QY 631 QEGLRKEPIHRVSAE 646
||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 233 KLLVKDPEKRLTAEE 248

RESULT 13
US-10-218-779-98
; Sequence 98, Application US/10218779
; Publication No. US2004002922A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsbrook II, John
```

APPLICANT: Lepley, Denise  
APPLICANT: Rieger, Daniel  
APPLICANT: Burgess, Catherine  
APPLICANT: Casman, Stacie  
APPLICANT: Spytek, Kimberly  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Mishra, Vishnu  
APPLICANT: Patturajan, Meera  
APPLICANT: Shenoy, Suresh  
APPLICANT: Rastelli, Luca  
APPLICANT: Tchernev, Velizar  
APPLICANT: Vernet, Corine  
APPLICANT: Zethusen, Bryan  
APPLICANT: Malyankar, Uriel  
APPLICANT: Guo, Xiaojia  
APPLICANT: Miller, Charles  
APPLICANT: Gangolli, Esha  
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-214  
CURRENT APPLICATION NUMBER: US/10/218,779  
CURRENT FILING DATE: 2002-08-14  
PRIOR APPLICATION NUMBER: 60/253,834  
PRIOR FILING DATE: 2000-11-29  
PRIOR APPLICATION NUMBER: 60/250,926  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 60/264,180  
PRIOR FILING DATE: 2001-01-25  
PRIOR APPLICATION NUMBER: 60/313,656  
PRIOR FILING DATE: 2001-08-20  
PRIOR APPLICATION NUMBER: 60/327,456  
PRIOR FILING DATE: 2001-10-05  
NUMBER OF SEQ ID NOS: 216  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 98  
LENGTH: 255  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-218-779-98

Query Match 7.3%; Score 367; DB 15; Length 255;

Best Local Similarity 37.9%; Pred. No. 1.4e-13;

Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

Qy 406 LGRSGFGEVHRMEDKOTGFCQAVKKVRLVFRAE-----ELMACAGLTSPRIVPLYGA 458  
Db 6 LGKAFGKVIYLRDKTKTKGLVAIKVIKKEKILKKRERILREIKILKKLDPNIVKLYDV 65  
Qy 459 VREGPWNIIFMELLEGGSLGQLVKEQGCLPEDRALYYLQALEGLELYLHRSRILHGDVKA 518  
Db 66 FEDDDKLYVMVEYCEGGDLFDLLKKRGLSEDEARFYARQILSALEYLHSGQIIHRDLKP 125  
Qy 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLITGDIYIPGTETHMAPEVLVGRSCDAKVD 578  
Db 126 ENILLSDSG-HVKLADFLGAKQLDSSG---TLLT--TFVGTPEYMAPEVLVGRGKGYKAVD 179  
Qy 579 VWSSCMMLHMLNGCHPWTQFFRG-----PLCLKIASPPPPVREIPP---SCAPLTAQAI 630  
Db 180 IWSGLVILYELTKPP-----FPGDDQLLALFKIKGKPPPP---FPPPEWKISPEAKDLI 232  
Qy 631 QEGLRKEPIHRVSAAE 646  
Db 233 KKLVLKQPEKELTAAE 248

RESULT 14

US-09-964-956-72

Sequence 72, Application US/09964956

Publication No. US20040043926A1

GENERAL INFORMATION:

APPLICANT: Gerlach, Valerie L

APPLICANT: MacDougall, John R

APPLICANT: Smithson, Glenda  
APPLICANT: Millet, Isabelle  
APPLICANT: Stone, David  
APPLICANT: Gunther, Erik  
APPLICANT: Ellerman, Karen  
APPLICANT: Grosse, William M  
APPLICANT: Alsbrook II, John P  
APPLICANT: Lepley, Denise M  
APPLICANT: Burgess, Catherine E  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Spytek, Kimberly A  
APPLICANT: Leach, Martin D  
APPLICANT: Shimkets, Richard A  
TITLE OF INVENTION: No. US20040043926A1el Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-124  
CURRENT APPLICATION NUMBER: US/09/964,956  
CURRENT FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/235,631  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/235,633  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/235,808  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,064  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,065  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,066  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/236,135  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: 60/237,434  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/238,321  
PRIOR FILING DATE: 2000-10-05  
PRIOR APPLICATION NUMBER: 60/238,399  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/238,396  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: 60/276,667  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/294,823  
PRIOR FILING DATE: 2001-05-31  
PRIOR APPLICATION NUMBER: 60/304,868  
PRIOR FILING DATE: 2001-07-12  
NUMBER OF SEQ ID NOS: 127  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 72  
LENGTH: 256  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:  
OTHER INFORMATION: Serine/threonine protein kinase Consensus  
OTHER INFORMATION: Sequence  
US-09-964-956-72

Query Match 7.3%; Score 367; DB 11; Length 256;

Best Local Similarity 37.9%; Pred. No. 1.4e-13;

Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;

Qy 406 LGRSGFGEVHRMEDKOTGFCQAVKKVRLVFRAS-----ELMACAGLTSPRIVPLYGA 458  
Db 7 LGKAFGKVIYLRDKTKTKGLVAIKVIKKEKILKKRERILREIKILKKLDPNIVKLYDV 66  
Qy 459 VREGPWNIIFMELLEGGSLGQLVKEQGCLPEDRALYYLQALEGLELYLHRSRILHGDVKA 518  
Db 67 FEDDDKLYVMVEYCEGGDLFDLLKKRGLSEDEARFYARQILSALEYLHSGQIIHRDLKP 126  
Qy 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLITGDIYIPGTETHMAPEVLVGRSCDAKVD 578  
Db 127 ENILLSDSG-HVKLADFLGAKQLDSSG---TLLT--TFVGTPEYMAPEVLVGRGKGYKAVD 180

QY 579 VWSGCCMLHMLNGCHPWTQFFRG-----PLCKIASPPPPVREIPP---SCAPLTAQAI 630  
Db 181 IWSGLVILYELLTGKPP-----FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233

QY 631 QEGLRKEPIHRVSAAE 646  
Db 234 KKLVLKDPKRLTAEE 249

RESULT 15  
US-10-055-569A-99  
; Sequence 99, Application US/10055569A  
; Publication No. US20040024181A1  
; GENERAL INFORMATION:  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Gilbert, Jennifer  
; APPLICANT: Casman, Stacie  
; APPLICANT: Bialock, Angela  
; APPLICANT: Li, Li  
; APPLICANT: Vernet, Corine  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Mishra, Vishnu S  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Gerlach, Valerie L  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Malyanker, Uriel  
; APPLICANT: Stone, David  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Gunther, Erik  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Anderson, David W  
; TITLE OF INVENTION: No. US20040024181A1el Human Proteins, Polynucleotides Encoding Th  
; FILE REFERENCE: 21402-191  
; CURRENT APPLICATION NUMBER: US/10/055,569A  
; PRIOR FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: 60/243,642  
; PRIOR FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 60/243,320  
; PRIOR FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 60/243,592  
; PRIOR FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 60/243,681  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: 60/243,863  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: 60/244,443  
; PRIOR FILING DATE: 2000-10-31  
; PRIOR APPLICATION NUMBER: 60/245,029  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/244,995  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/245,293  
; PRIOR FILING DATE: 2000-11-02  
; PRIOR APPLICATION NUMBER: 60/245,315  
; PRIOR FILING DATE: 2000-11-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 99  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-055-569A-99

QY 406 LGRSGFGEVHRMEDKQTFQCAVKKVRLEVFRAE-----ELMACAGLTSPRIVPLYGA 458  
Db 7 LGKGAFAKVIYARDKKTGKLVAKVIKKEKLKKKRERILREIKILKKLDHPNIVKLYDV 66

QY 459 VREGFWYNIIFMELLEGGSLGOLVKBOCCLPEDRALYVLGQALEGLELYLHRRILHGDVKA 518  
Db 67 FEDDDKLYLVMEYCEGGDLFDLLKKRGLSEDEARFYARQILSALEYLHSGQIHRDLKP 126

QY 519 DNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIYIPGTETHMAPEVVLGRSCDAKVD 578  
Db 127 ENILLSDG-HVKLADFLAKQLDSDG---TLLT--TFVGTPEYMAPEVLLGKGYGKAVD 180

QY 579 VWSGCCMLHMLNGCHPWTQFFRG-----PLCKIASPPPPVREIPP---SCAPLTAQAI 630  
Db 181 IWSGLVILYELLTGKPP-----FPGDDQLLALFKKIGKPPPP---FPPPEWKISPEAKDLI 233

QY 631 QEGLRKEPIHRVSAAE 646  
Db 234 KKLVLKDPKRLTAEE 249

Search completed: June 7, 2005, 12:25:35  
Job time : 121.819 secs

Query Match 7.3%; Score 367; DB 15; Length 256;  
Best Local Similarity 37.9%; Pred. No. 1.4e-13;  
Matches 97; Conservative 37; Mismatches 94; Indels 28; Gaps 8;



Query Match	43.1%	Score 1133.4	DB 2	Length 3156
Best Local Similarity	94.8%	Pred. No. 0		
Matches 1241	Conservative 4	Mismatches 43	Indels 21	Gaps 7
Qy 1	CCCTCTCAGAGCCCAAGCCCATCAAGAGGGGTGAGGAAAGCCCATCCACCGCGTGT	60		
Db 1868	CCCTCTCTCAGAGCCCAAGCCCATCAAGAGGGGTGAGGAAAGCCCATCCACCGCGTGT	1927		
Qy 61	CTGCAGCCGAGCTGGGAGGGAAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTTCTGAAGA	120		

Db 1928 CTGAGCGAGCTGGGAGGAGGTGAACCGGCGCATACAGCAAGTGGGAGTCTGAAGA 1987  
QY 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCAAAATCAAGCCAAATTACC 180  
Db 1988 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCAAAATCAAGCCAAATTACC 2047  
QY 181 ACCAGACCTTCATCCCGAGCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGGCAG 240  
Db 2048 ACCAGACCTTCATCCCGAGCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGGCAG 2107  
QY 241 CTGAGGAGACAACAGCGAGAGCCCTTAAGCTTCAGCTTCCTCTCCACAGAGCCCGCAG 300  
Db 2108 CTGAGGAGACAACAGCGAGAGCCCTTAAGCTTCAGCTTCCTCTCCACAGAGCCCGCAG 2167  
QY 301 AGCCAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCCCT 360  
Db 2168 AGCCAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGGAACCCCT 2227  
QY 361 TACCTCTGTCTCTCTGGAGCAGCCCTGCGCAGAAACCCAGCTTCACAGAGCGGAAAG 420  
Db 2228 TACCTCTGTCTCTCTGGAGCAGCCCTGCGCAGAAACCCAGCTTCACAGAGCGGAAAG 2287  
QY 421 CAACCGTCCCGAGCAGGAGTCTGAGCAGCTGGAATAGAAATTTCTCTCAACAGCCTGT 480  
Db 2288 CAACCGTCCCGAGCAGGAGTCTGAGCAGCTGGAATAGAAATTTCTCTCAACAGCCTGT 2347  
QY 481 CCAGCATTCTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACGCC 540  
Db 2348 CCAGCATTCTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACGCC 2407  
QY 541 TCTCCCTGTCTGATGACAGTGAAGAAACCCATCAAGAGCTCTCAAGCTTCGCGGGACA 600  
Db 2408 TCTCCCTGTCTGATGACAGTGAAGAAACCCATCAAGAGCTCTCAAGCTTCGCGGGACA 2467  
QY 601 CCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCGGAGCTCGAAGCTTCAGCTCGA 660  
Db 2468 CCTGAGCTCAGCGTACACTCTCTGAGCAGCAGCGGAGCTCGAAGCTTCAGCTCGA 2527  
QY 661 ACATGTGTCTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCATATGTGTGAAG 720  
Db 2528 ACATGTGTCTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCATATGTGTGAAG 2587  
QY 721 TCAAAATACAGTCTCTTAATGTGAACACTGACATCCGGAGTTCACCGGGTCAAG 780  
Db 2588 TCAAAATACAGTCTCTTAATGTGAACACTGACATCCGGAGTTCACCGGGTCAAG 2647  
QY 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTCAGCTTGCTCA 840  
Db 2648 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTCAGCTTGCTCA 2707  
QY 841 CCAAGACGGGAGCCTCTTTCGTACGACATGAGGTGCAGACTCGGGGATCGACCTGC 900  
Db 2708 CCAAGACGGGAGCCTCTTTCGTACGACATGAGGTGCAGACTCGGGGATCGACCTGC 2767  
QY 901 AGTGACACTGGCCCTGATGGCAGCTTCGCTGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 2768 AGTGACACTGGCCCTGATGGCAGCTTCGCTGAGCTGGAGGTCAAGCATGGCCAGC 2827  
QY 961 TGGAGAACAGGCGCTTAACCTCTGCTCCACCGCGGCTCCACACTCTCCGGAAGCAGCCT 1020  
Db 2828 TGGAGAACAGGCGCTTAACCTCTGCTCCACCGCGGCTCCACACTCTCCGGAAGCAGCCT 2887  
QY 1021 TCTGTCTGGTGACAGATGTGCTGCTGAACACAGCTCAGCGCTTCAGGGGATYTG 1080  
Db 2888 TCTGTCTGGTGACAGATGTGCTGCTGAACACAGCTCAGCGCTTCAGGGGATYTG 2945  
QY 1081 NCCAGCCCCCGGCTCAGAGTGGGAACAGGGCTTCGNCAGCAGCAGAGGAGGAGG 1140  
Db 2946 GCGAGCCCCCGGCTCAGAGTGGGAACAGGGCTTCGNCAGCAGCAGAGGAGGAGG 2998  
QY 1141 CAAGCAGAAATGCTCCAGGATTTACANCTGTAGCCGCTGAGCCGCTCCGAGGAG 1200

Db 2999 GCAAGCAGAAATGCTCCAGGATTTTCAACACTGAGCCCTGCCCCAC-----CCTGCTGA 3052  
QY 1201 AAACATVTCGCCACCTGTAAGAGACAGAGGAGGATGNCAGAGTTNNACCTYGGGGAA 1260  
Db 3053 AAAAATCATCCGCCACCTGTAAGAGACAGAGGAGGATGGC-----AGGAGTTACCTGGGAA 3108  
QY 1261 ACAAACACAGGATCTTTTCTGCTCCCTGCTCCAGTNCAGATTTGGGCTG 1309  
Db 3109 ACAAACACAGGATCTTTTCTGCTCCCTGCTCCAGTNCAGATTTGGGCTG 3155  
  
RESULT 2  
US-09-023-321-1  
; Sequence 1, Application US/09023321  
; Patent No. 5844073  
; GENERAL INFORMATION:  
; APPLICANT: Rothe, Mike  
; APPLICANT: Wu, Lin  
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,321  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,518  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: T97-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3156 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-09-023-321-1  
  
Query Match 43.1%; Score 1133.4; DB 2; Length 3156;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;  
  
QY 1 CCCCTCTCACAGCCAGCCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCTGT 60  
Db 1868 CCCCTCTCACAGCCAGCCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCTGT 1927  
QY 61 CTGCAAGCGAGCTGGGAGGAGGTGAACCCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
Db 1928 CTGCAAGCGAGCTGGGAGGAGGTGAACCCGGGCACTACAGCAAGTGGAGGTCTGAAGA 1987  
QY 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCAAAATCAAGCCAAATTACC 180  
Db 1988 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCAAAATCAAGCCAAATTACC 2047  
QY 181 ACCAGACCTTCATCCCGAGCGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGGCAG 240

Db 2048 ACAGACCTCCATGCCAGCGGAGAGAGCTTTTCGCCAAGGGCCCAAGGGCCCGGCCAG 2107  
Qy 241 CTGAGGAGACAACAGGAGAGCCCTTAAGCTTCCAGCCCTCTCTCCCAACAGAGCCCCCAG 300  
Db 2108 CTGAGGAGACAACAGGAGAGCCCTTAAGCTTCCAGCCCTCTCTCCCAACAGAGCCCCCAG 2167  
Qy 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 360  
Db 2168 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 2227  
Qy 361 TACCTCTGTCTCTCCCTGAGCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAG 420  
Db 2228 TACCTCTGTCTCTCCCTGAGCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAG 2287  
Qy 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATATTAATCTCTCAACAGCCCTGT 480  
Db 2288 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATATTAATCTCTCAACAGCCCTGT 2347  
Qy 481 CCCAGCATTCTCTGAGGAGCAGGAGCAAAATCTCTGTGCTCAGCATCGACAGCC 540  
Db 2348 CCCAGCATTCTCTGAGGAGCAGGAGCAAAATCTCTGTGCTCAGCATCGACAGCC 2407  
Qy 541 TCTCCCTGTGGATGACAGTGAAGAACCCATCAAGAGCCCTCTCAAGCTTCGGGACA 600  
Db 2408 TCTCCCTGTGGATGACAGTGAAGAACCCATCAAGAGCCCTCTCAAGCTTCGGGACA 2467  
Qy 601 CCCTGAGCTCAGGCGTACACTCTGAGCAGCAGGCGGAGCTTCAAGCTTCCAGCTGGA 660  
Db 2468 CCCTGAGCTCAGGCGTACACTCTGAGCAGCAGGCGGAGCTTCAAGCTTCCAGCTGGA 2527  
Qy 661 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720  
Db 2528 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2587  
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACGGGTCAAG 780  
Db 2588 TCCAAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCCACGGGTCAAG 2647  
Qy 781 TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTTGCAGCTTCAAGCTTGTGCA 840  
Db 2648 TGGGAGACATCGCACTGGCATCAGCAGCAGATCCAGCTTGCAGCTTCAAGCTTGTGCA 2707  
Qy 841 CCAAAGACGGCAGCCTGTTCCGTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 900  
Db 2708 CCAAAGACGGCAGCCTGTTCCGTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 2767  
Qy 901 AGTGCACATCGGCCCTGATGGAGCTTGGCTGGAGCTGAGGCTCAAGATGGCCAGC 960  
Db 2768 AGTGCACATCGGCCCTGATGGAGCTTGGCTGGAGCTGAGGCTCAAGATGGCCAGC 2827  
Qy 961 TGGAGAACAGCCCTTAACCTCGCCCTCCACCGCGGCTTCCACACTGCGGAAAGCAGCCT 1020  
Db 2828 TGGAGAACAGCCCTTAACCTCGCCCTCCACCGCGGCTTCCACACTGCGGAAAGCAGCCT 2887  
Qy 1021 TCCTGTCTGGTGACGATGCTGCTGAAACACAGGCTCAGCCGCTTCCAGGGGATYTG 1080  
Db 2888 TCCTGTCTGGTGACGATGCTGCTGAAACACAGGCTCAGCCGCTTCCAGGGGATYTG 2945  
Qy 1081 NCCAGCCCCCGGTTCACAGTGGGAACAGGCGCTTCGNAGCNAGNAGTNGGGGG 1140  
Db 2946 GCCAGCCCCCGGTTCACAGTGGGAACAGGCGCTTCG-----CAGCAGCAAGGTGGG 2998  
Qy 1141 CAAGNAGATGCTCCAGATTTTCACTGAGCCCTGCCCCCAGCCCTGCTGAADA 1200  
Db 2999 GCAAGCAGAAATGCTCCAGATTTTCACTGAGCCCTGCCCCCAGCCCTGCTGA 3052  
Qy 1201 AAACATYTCGCCAGTGAAGAGACAGAGGAGATGNCAGGATTTNACCTYGGGGAA 1260  
Db 3053 AAACATYTCGCCAGTGAAGAGACAGAGGAGATGNCAGGATTTNACCTYGGGGAA 3108  
Qy 1261 ACAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTGTGGCCTG 1309  
Db 3109 ACAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTGTGGCCTG 3155

RESULT 3  
US-09-032-475-1  
; Sequence 1, Application US/09032475  
; Patent No. 5854003  
; GENERAL INFORMATION:  
; APPLICANT: Roche, Mike  
; APPLICANT: Wu, Lin  
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032,475  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/887,518  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: T97-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3156 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-032-475-1

Query Match 43.1%; Score 1133.4; DB 2; Length 3156;  
Best Local Similarity 94.8%; Pred. No. 0;  
Matches 1241; Conservative 4; Mismatches 43; Indels 21; Gaps 7;  
Qy 1 CCCCTCTCACAGCCCGGAGGCTTCAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60  
Db 1868 CCCCTCTCACAGCCCGGAGGCTTCAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 1927  
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGGCTTACAGCAAGTGGGAGGTCTGAAGA 120  
Db 1928 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGGCTTACAGCAAGTGGGAGGTCTGAAGA 1987  
Qy 121 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 180  
Db 1988 GCCCTTGGAGGGAGAGATATAAGAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 2047  
Qy 181 ACCAGACCCCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGGGCCCGGCCAG 240  
Db 2048 ACCAGACCCCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGGGCCCGGCCAG 2107  
Qy 241 CTGAGGAGACAACAGGCGAGAGCCCTTAAGCTTCCAGCCCTCTCTCCCAACAGAGCCCCCAG 300  
Db 2108 CTGAGGAGACAACAGGCGAGAGCCCTTAAGCTTCCAGCCCTCTCTCCCAACAGAGCCCCCAG 2167  
Qy 301 AGCCAAACAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 360  
Db 2168 AGCCAAACAAGTCTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGATGTGGAAACCT 2227

```
QY 361 TACCTCTGTCTCTCTGGAGCAGCCCTGCCAGAAACCCAGCTCACCAGACGGGAAG 420
Db 2228 TACCTCTGTCTCTCTCTGGAGCAGCCCTGCCAGAAACCCAGCTCACCAGACGGGAAG 2287
QY 421 CAACCGTCCGGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATTCCTCAACAGCCTGT 480
Db 2288 CAACCGTCCGGAGCAGGAACCTGCAGCAGCTGGAATAGAAATATTCCTCAACAGCCTGT 2347
QY 481 CCAGCAGCTTTCTCTGGAGGAGCAGGAGCAATCTCTCGTGCCTCAGCATCGACAGCC 540
Db 2348 CCAGCAGCTTTCTCTGGAGGAGCAGGAGCAATCTCTCGTGCCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTCTGGATGACAGTGAAGAACCCATCAAGAGCCTCTCAAGCTCGCGGACA 600
Db 2408 TCTCCCTGTCTGGATGACAGTGAAGAACCCATCAAGAGCCTCTCAAGCTCGCGGACA 2467
QY 601 CCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCGCAGGCTCGAAGCTCAGCTGGA 660
Db 2468 CCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCGCAGGCTCGAAGCTCAGCTGGA 2527
QY 661 ACATGCTGTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2528 ACATGCTGTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACTCTGCACATCCGGGAGTTCCACCGGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAACACTCTGCACATCCGGGAGTTCCACCGGGTCAAAG 2647
QY 781 TGGGAGACATCGGCATCGCATCAGGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
Db 2648 TGGGAGACATCGGCATCGCATCAGGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 2707
QY 841 CAAAGAGCGGAGCTGTTCCTCTAGCATGAGTGGTGCAGCTCGGCGCATCGACCTGC 900
Db 2708 CAAAGAGCGGAGCTGTTCCTCTAGCATGAGTGGTGCAGCTCGGCGCATCGACCTGC 2767
QY 901 AGTGACACTGGCCCTCTGATGGCAGCTTCGCCCTGGAGCTGGAGGCTCAAGCATGGCCAGC 960
Db 2768 AGTGACACTGGCCCTCTGATGGCAGCTTCGCCCTGGAGCTGGAGGCTCAAGCATGGCCAGC 2827
QY 961 TGGGAAACAGGCGCTTAACTCTGCCCTCAACCGCGGCTCCACACTGCGGGAAGCAGCT 1020
Db 2828 TGGGAAACAGGCGCTTAACTCTGCCCTCAACCGCGGCTCCACACTGCGGGAAGCAGCT 2887
QY 1021 TCTCTGCTGCTGACAGTCTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG 1080
Db 2888 TCTCTGCTGCTGACAGTCTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG 2945
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACAGGGCTCTGNCAGCNAGNAGGNTGGGGG 1140
Db 2946 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGCTCG----CAGCAGCAAGGTGGGG 2998
QY 1141 CAAGCNAGATGCTCCAGGATTTTCANCTGAGCCNTGCCANCCCTGCTGAADA 1200
Db 2999 GCAAGCAGATGCTCCAGGATTTTCACCTGAGCCCTGCCCCAC-----CCTGTGTA 3052
QY 1201 AAACATVTCGCGCAGTGAAGAGACAGAGGAGTGNACAGGATTTNACCTYGGGAA 1260
Db 3053 AAACATVTCGCGCAGTGAAGAGACAGAGGAGTGNACAGGATTTNACCTYGGGAA 3108
QY 1261 ACAAACAGGATCTTTTCTGCCCCCTGCTCAGTNCAGTGGGCTG 1309
Db 3109 ACAAACAGGATCTTTTCTGCCCCCTGCTCAGTNCAGTGGGCTG 3155
```

RESULT 4

```
US-09-257-703-2
; Sequence 2, Application US/09257703
; Patent No. 626538
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
```

```
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
; FILE REFERENCE: 30448.61USU1
; CURRENT APPLICATION NUMBER: US/09/257,703
; EARLIER FILING DATE: 1999-02-25
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; TYPE: DNA
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-257-703-2
```

```
Query Match 37.1%; Score 977; DB 3; Length 2844;
Best Local Similarity 100.0%; Pred. No. 6.3e-272;
Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CCCCTCTCACAGCCAGCGCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1868 CCCCTCTCACAGCCAGCGCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 1927
QY 61 CTGACAGCGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120
Db 1928 CTGACAGCGAGCTGGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 1987
QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 1988 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2047
QY 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCAAGGGGCCCGGGCCCGGCGCAG 240
Db 2048 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTCGCAAGGGGCCCGGGCCCGGCGCAG 2107
QY 241 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTTCCAGCCTCTCTCCACAGAGCCCCCAG 300
Db 2108 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTTCCAGCCTCTCTCCACAGAGCCCCCAG 2167
QY 301 AGCCAAACAAGTCTCTCCCTTGAATTTTGACAGAGGAGAGTCTGGGATGTGGGAACCCCT 360
Db 2168 AGCCAAACAAGTCTCTCCCTTGAATTTTGACAGAGGAGAGTCTGGGATGTGGGAACCCCT 2227
QY 361 TACTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAG 420
Db 2228 TACTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGAAG 2287
QY 421 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATATTAATTCCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATATTAATTCCTCAACAGCCTGT 2347
QY 481 CCGAGCCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGTCCTCAGCATCGACAGCC 540
Db 2348 CCGAGCCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGTCCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAGGGCCTCTCAAGCTCGCGGGAACA 600
Db 2408 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAGGGCCTCTCAAGCTCGCGGGAACA 2467
QY 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTCAGCTGGA 660
Db 2468 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTCAGCTGGA 2527
QY 661 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720
Db 2528 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACTCTGCACATCCCGGGAGTTCCACCGGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAACACTCTGCACATCCCGGGAGTTCCACCGGGTCAAAG 2647
QY 781 TGGGAGACATCGGCATCGCATCAGGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
Db 2887 TGGGAGACATCGGCATCGCATCAGGAGCAGATCCAGCTCGAGCTTCAGCTTGGTCA 840
```

Db 2648 TGGGAGACATCCCACTGGCATCAGCAGATCCCAAGCTTCCAGCTTTCAGCTTGGTCA 2707  
Qy 841 CCAAAGACGGGAGCGCTTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900  
Db 2708 CCAAAGACGGGAGCGCTTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 2767  
Qy 901 AGTGACACTGGCGCCCTGATGGAGCTTCGGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 2768 AGTGACACTGGCGCCCTGATGGAGCTTCGGCTGGAGCTGGAGGTCAAGCATGGCCAGC 2827  
Qy 961 TGGAGAACAGGCCCTAA 977  
Db 2828 TGGAGAACAGGCCCTAA 2844

## RESULT 5

US-09-871-889A-2

; Sequence 2, Application US/09871889A

; Patent No. 6645728

; GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

; APPLICANT: Gelezuinas, Romas

; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF- $\alpha$ 

; FILE OF INVENTION: IL-1

; FILE REFERENCE: 30448.61USDI

; CURRENT APPLICATION NUMBER: US/09/871,889A

; CURRENT FILING DATE: 2001-06-01

; PRIOR APPLICATION NUMBER: 09/257,703

; PRIOR FILING DATE: 1999-02-25

; PRIOR APPLICATION NUMBER: 60/076,299

; PRIOR FILING DATE: 1998-02-27

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 2844

; TYPE: DNA

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-871-889A-2

Query Match 37.1%; Score 977; DB 4; Length 2844;

Best Local Similarity 100.0%; Pred. No. 6.3e-272;

Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCCTCTCACAGCCAGCCATCCAGAGGGGCTGAGAAAGAGCCATCCACCGCTGT 60  
Db 1868 CCCCTCTCACAGCCAGCCATCCAGAGGGGCTGAGAAAGAGCCATCCACCGCTGT 1927  
Qy 61 CTGACGGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
Db 1928 CTGACGGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 1987  
Qy 121 GCCCTTGGAGGGAGATATTAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC 180  
Db 1988 GCCCTTGGAGGGAGATATTAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC 2047  
Qy 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGGCCCGCCAG 240  
Db 2048 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGGCCCGCCAG 2107  
Qy 241 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAG 300  
Db 2108 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAG 2167  
Qy 301 AGCCAAACAAGTCTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 360  
Db 2168 AGCCAAACAAGTCTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCT 2227  
Qy 361 TACCTCTGTCTCTCTGGAGCCAGCCCTGCGCAAGAAACCCAGCTCAACGAGCGGAAG 420  
Db 2228 TACCTCTGTCTCTCTGGAGCCAGCCCTGCGCAAGAAACCCAGCTCAACGAGCGGAAG 2287

Qy 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT 480  
Db 2288 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAATAGAAATTAATTCCTCAACAGCCTGT 2347  
Qy 481 CCAGAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCCAGAGCC 540  
Db 2348 CCAGAGCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCCAGAGCC 2407  
Qy 541 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAAAGGCTCTTCAAGCTTCGCGGGACA 600  
Db 2408 TCTCCCTGTCCGATGACAGTGAAGAACCCATCAAAAGGCTCTTCAAGCTTCGCGGGACA 2467  
Qy 601 CCCTGAGCTCAGCGGTACACTCTGAGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTGGA 660  
Db 2468 CCCTGAGCTCAGCGGTACACTCTGAGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTGGA 2527  
Qy 661 ACATGGTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 720  
Db 2528 ACATGGTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 2587  
Qy 721 TCCAAATACAGTCTCTTAATGGTGAAACCTTCGACATCCCGGGAGTTCACCGGCTCAAG 780  
Db 2588 TCCAAATACAGTCTCTTAATGGTGAAACCTTCGACATCCCGGGAGTTCACCGGCTCAAG 2647  
Qy 781 TGGGAGACATCGGCACATGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA 840  
Db 2648 TGGGAGACATCGGCACATGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA 2707  
Qy 841 CCAAAGACGGGAGCGCTTTCGCTACGACATGGAGGTGCCAGATCCGGGATCCAGCTGC 900  
Db 2708 CCAAAGACGGGAGCGCTTTCGCTACGACATGGAGGTGCCAGATCCGGGATCCAGCTGC 2767  
Qy 901 AGTGACACTGGCCCTCGATGGCAGCTTTCGCTGGAGCTTGAGGCTCAAGCATGGCCAGC 960  
Db 2768 AGTGACACTGGCCCTCGATGGCAGCTTTCGCTGGAGCTTGAGGCTCAAGCATGGCCAGC 2827  
Qy 961 TGGAGAACAGGCCCTAA 977  
Db 2828 TGGAGAACAGGCCCTAA 2844

## RESULT 6

US-09-252-991A-9023

; Sequence 9023, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 9023

; LENGTH: 825

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9023

Query Match 1.6%; Score 42.4; DB 4; Length 825;

Best Local Similarity 54.5%; Pred. No. 0.099;

Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 744 GAACACTGCACATCCGGAGTTCCACCGGTTCAAAAGTGGAGACATCGCCACTGGCATC 803  
Db 659 GAACACATGGCCAAAGCAGCCCTGCAGCATGGCAAGCCGATGAAGATCGCCATTCGGC 718  
Qy 804 AGCAGCCAGATCCAGCTGCGAGCTTTCAGCTTGGTTCACCAAAAGACGGGAGCTTTCGC 863

Db 719 GCGGGCAGAGCGCGCGGAGCCTTCATCGACCTCAACGACAGCTACCCGTCGGTGCAG 778

QY 864 TAGCATGAGGTGCCAGACTCGGCGCATCGACCTG 899

Db 779 GCGGACATGATCTCGGTGCTCGGCGCTCAAGCG 814

## RESULT 7

US-09-252-991A-9108

; Sequence 9108, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 9108

; LENGTH: 1359

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-9108

Query Match 1.6%; Score 42.4; DB 4; Length 1359;

Best Local Similarity 54.5%; Pred. No. 0.13;

Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 744 GAACACCTGCACATCGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATC 803

Db 610 GAACACATGGCCAAAGCAGCCCTGCGAGCAGTGGCAAGCGGATGAAGATCGCCATTATCGGC 669

QY 804 AGCAGCCAGATCCAGCTCGACCTTCAGCTTGTGCACCAAGACGGCGACCTGTCGC 863

Db 670 GCGCGCAGAGCGCGGAGGCTTCATCGACCTCAACGACAGCTACCCGTCGGTGCAG 729

QY 864 TAGCATGAGGTGCCAGACTCGGCGCATCGACCTG 899

Db 730 GCGGACATGATCTCGTGCCTCGCGCTCAAGCG 765

## RESULT 8

US-09-252-991A-8985/c

; Sequence 8985, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 8985

; LENGTH: 1746

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-8985

Query Match 1.6%; Score 42.4; DB 4; Length 1746;

Best Local Similarity 54.5%; Pred. No. 0.15;

Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 744 GAACACCTGCACATCGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATC 803

Db 438 GAACACATGGCCAAGCAGCCCTGCAGCAGTGGCAAGCGCATGAAGATCGCCATTATCGGC 379

QY 804 AGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGTGTACCAAGAGCGGCGACGCTGTTGCG 863

Db 378 GCGGGCAGAGCGCGGAGGCTTCATCGACCTCAACGACAGCTACCCGTCGGTGCAG 319

QY 864 TAGCATGAGGTGCCAGACTCGGCGCATCGACCTG 899

Db 318 GCGGACATGATCTCGGTGCTCGGCGCTCAAGCG 283

## RESULT 9

5223425-7

; Patent No. 5223425

; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

; BARRY M.; WHITE, TYLER R.

; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY

; NUMBER OF SEQUENCES: 19

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/277,963

; FILING DATE: 30-NOV-1988

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 34,203

; FILING DATE: 02-APR-1987

; SEQ ID NO:7

; LENGTH: 835

5223425-7

Query Match 1.5%; Score 40.2; DB 6; Length 835;

Best Local Similarity 48.1%; Pred. No. 0.43;

Matches 114; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 747 CACTGCACATCCGGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGC 806

Db 494 CGCCCGGACAGCTGCGACGCTCTTTCAGTGTGGACCGCGCCACCTGCAACCGG 553

QY 807 AGCCAGATCCAGCTCGAGCCTTCAGCTTGTACCAAGACGGCGAGCCTGTTGCTAC 866

Db 554 CGCAGCACACGACGCGGCCATCCGAGCGCTTGATGTCGGGAGAGCAATCGCGG 613

QY 867 GACATGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCTTGATGGGAGC 926

Db 614 GACAGCTGCAAGGTGACTCGGGGCGCGCTGTGTGCGGGGCGTCTCGAGGCGTG 673

QY 927 TTCGCTGGAGTCAAGCATGGCAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983

Db 674 GTCACCTCGGCTCGCGGTTTTCGGGCAACCGCAAGAGCCCGGATCTACACCGCG 730

## RESULT 10

5223425-7

; Patent No. 5223425

; APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

; BARRY M.; WHITE, TYLER R.

; TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

; D ACTIVITY

; NUMBER OF SEQUENCES: 19

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/277,963

; FILING DATE: 30-NOV-1988

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 34,203

; FILING DATE: 02-APR-1987

; SEQ ID NO:7

; LENGTH: 835

5223425-7

Query Match

Best Local Similarity 48.1%; Pred. No. 0.43; Length 835;

Matches 114; Conservative 0; Mismatches 123; Indels 0; Gaps 0;



```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,284
; FILING DATE: 02-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/570,184
; FILING DATE: 08-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1506/CIP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1005 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; STRAIN: single
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1005
; OTHER INFORMATION: /product= "alternate sigma factor
; OTHER INFORMATION: RpoS"
;
; US-08-287-442-8
;
Query Match 1.5%; Score 39.6; DB 1; Length 1005;
Best Local Similarity 46.5%; Pred.No. 0.72; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 145;

QY 783 GGAGACATCCCACTGGCATCAGACGACAGATCCCAGTCGACGCTTCAGCTTGGTCAACC 842
Db 601 GNAGAAATCCCAACCTGTCTGGAGAAACCGTAGGTGAGGTCAAGCGCATGCTGGGTCTC 660
QY 843 AAGACGGGAGCGCTGTTCGCTAGCATGAGGTGCCAGACTCGGSCATCGACCTGCAG 902
Db 661 AATGAGCGGGTGCTTCAGTCGACGCTCTCGCTGGGTCGGGATTCGGNATAAACCCCTG 720
QY 903 TGCACACTGGCCCTCGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTG 962
Db 721 GACACCTTCACCGACGATCGCCAAACCGATCCGTCGAGCTGCTGCAGGATGACGACCTG 780
QY 963 GAGAAACAGGCCCTAACCCCTGGCCCTCCACCGCCGGCTCCACACTGCCGGAAGACGCTTC 1022
Db 781 TCGCAAAAGCATCGATCATGTGCTTTCGAACTGACCGACAGCGGTGAAGTAGTGGTT 840
QY 1023 CTGCTCGGTGCACGATGCTGCCCTGAAAAACA 1053
Db 841 CGCCGCTTCGGCTTGGCGGCCATGAAAGCA 871

Search completed: June 10, 2005, 20:46:45
Job time : 303.303 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 999.861 Seconds  
(without alignments)

16311.995 Million cell updates/sec

Title: US-09-155-676B-3

Perfect score: 2631

Sequence: 1 ccccttcacagccagcc.....accctgacctcagggggg 2631

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT NEW PUB.seq.\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW PUB.seq.\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW PUB.seq.\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW PUB.seq.\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW PUB.seq.\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
  - 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
  - 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
  - 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
  - 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
  - 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW PUB.seq.\*
  - 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
  - 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW PUB.seq.\*
  - 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW PUB.seq.\*
  - 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1850.2	70.3	4596	10	US-09-981-397A-17
2	1850.2	70.3	4596	17	Sequence 17, Appl
3	1850.2	70.3	4596	18	Sequence 975, App
4	1850.2	70.3	4596	19	Sequence 975, App
5	1783.4	67.8	4684	13	US-10-087-192-887
6	977	37.1	2844	9	US-09-871-889-2
7	973.4	37.0	73967	13	US-10-087-192-886

8	754.2	28.7	1385	18	US-10-296-115-97	Sequence 97, Appl
9	672.8	25.6	2828	13	US-10-087-192-884	Sequence 884, App
c 10	349.4	13.3	2475	17	US-10-108-260A-1754	Sequence 1754, Ap
11	242.2	9.2	499	10	US-09-918-995-2585	Sequence 2585, Ap
12	193	7.3	296	9	US-09-783-590-11852	Sequence 11852, A
13	188.8	7.2	42586	13	US-10-087-192-883	Sequence 883, App
c 14	88.4	3.4	455	20	US-10-357-930-22052	Sequence 22052, A
c 15	88.4	3.4	455	20	US-10-357-930-27909	Sequence 27909, A
c 16	88.4	3.4	476	20	US-10-357-930-12325	Sequence 12325, A
c 17	88.4	3.4	504	20	US-10-357-930-33496	Sequence 33496, A
c 18	88.4	3.4	504	20	US-10-357-930-42420	Sequence 42420, A
c 19	73.8	2.8	463	20	US-10-357-930-31556	Sequence 31556, Ap
c 20	52.8	2.0	65	10	US-09-908-975-25736	Sequence 25736, A
21	50	1.9	50	17	US-10-131-827-2583	Sequence 2583, Ap
22	50	1.9	50	17	US-10-131-827-3841	Sequence 3841, Ap
23	49	1.9	60	10	US-09-908-975-19057	Sequence 19057, A
24	43	1.6	447	20	US-10-425-115-153036	Sequence 153036, A
25	42.6	1.6	1464	20	US-10-258-148-11	Sequence 11, Appl
c 26	42.4	1.6	867	14	US-10-123-153-20	Sequence 20, Appl
c 27	42.4	1.6	867	15	US-10-156-761-5737	Sequence 5737, Ap
c 28	42.4	1.6	867	15	US-10-146-731-20	Sequence 20, Appl
c 29	42.4	1.6	867	15	US-10-140-472-20	Sequence 20, Appl
c 30	42.4	1.6	867	15	US-10-141-761-20	Sequence 20, Appl
c 31	42.4	1.6	867	16	US-10-142-885-20	Sequence 20, Appl
c 32	42.4	1.6	867	16	US-10-158-790-20	Sequence 20, Appl
c 33	42.4	1.6	867	17	US-10-137-871-20	Sequence 20, Appl
c 34	42.4	1.6	867	17	US-10-140-923-20	Sequence 20, Appl
c 35	42.4	1.6	867	17	US-10-141-756-20	Sequence 20, Appl
c 36	42.4	1.6	867	17	US-10-141-759-20	Sequence 20, Appl
c 37	42.4	1.6	867	17	US-10-140-803-20	Sequence 20, Appl
c 38	42.4	1.6	867	17	US-10-140-864-20	Sequence 20, Appl
c 39	42.4	1.6	867	18	US-10-142-426-20	Sequence 20, Appl
c 40	42.4	1.6	9025608	15	US-10-156-761-1	Sequence 1, Appl1
c 41	40.4	1.5	1089	17	US-10-282-122A-14856	Sequence 14856, A
c 42	40.2	1.5	639	9	US-09-821-255-1	Sequence 1, Appli
c 43	40.2	1.5	762	14	US-10-202-676-3	Sequence 3, Appli
c 44	40.2	1.5	835	14	US-10-202-676-1	Sequence 1, Appli
c 45	40.2	1.5	911	13	US-10-044-090-214	Sequence 214, App

ALIGNMENTS

RESULT 1

US-09-981-397A-17  
; Sequence 17, Application US/09981397A  
; Publication No. US20030082519A1  
; GENERAL INFORMATION:  
; APPLICANT: Axxima Pharmaceuticals AG  
; APPLICANT: Schubart, Daniel  
; APPLICANT: Habenberger, Peter  
; APPLICANT: Stein-Geisler, Matthias  
; APPLICANT: Bevec, Dorian  
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their  
; FILE REFERENCE: AXM-004.1 US  
; CURRENT APPLICATION NUMBER: US/09/981,397A  
; PRIOR FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/240,750  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-981-397A-17

Query Match 70.3%; Score 1850.2; DB 10; Length 4596;  
Best Local Similarity 90.3%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;  
QY 1 CCCCTCTCACAGCCAGCCATCCAGAGGGGTGAGGAAGAGCCATCCACCGCGTGT 60

Db 2100 CCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAGAGCCCATCCACCGGTGT 2159  
QY 61 CTGACGGAGCTGGAGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
Db 2160 CTGACGGAGCTGGAGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219  
QY 121 GCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCAATTACC 180  
Db 2220 GCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCAATTACC 2279  
QY 181 ACCAGACCTTCATCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCGGCCAG 240  
Db 2280 ACCAGACCTTCATCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCGGCCAG 2339  
QY 241 CTGAGAGACAACAGAGAGAGCCCTAAGCTCCAGCCCTCTCTCCACAGAGCCCCAG 300  
Db 2340 CTGAGAGACAACAGAGAGAGCCCTAAGCTCCAGCCCTCTCTCCACAGAGCCCCAG 2399  
QY 301 AGCCAAAACAAGTCTCTCTCTTGAATTTGAGCAAGAGAGTCTGGGATGTGGGAACCT 360  
Db 2400 AGCCAAAACAAGTCTCTCTCTTGAATTTGAGCAAGAGAGTCTGGGATGTGGGAACCT 2459  
QY 361 TACCTCTGTCTCTCTCTGGAGCCAGCCCTTGCCAGAAAACCCAGCTTCAACAGAGCGGAAAG 420  
Db 2460 TACCTCTGTCTCTCTCTGGAGCCAGCCCTTGCCAGAAAACCCAGCTTCAACAGAGCGGAAAG 2519  
QY 421 CAACCGTCCCGAGCAGGNACTGCAGCAGCTGGAATAGAAATATTTCTCTCAACAGCCGT 480  
Db 2520 CAACCGTCCCGAGCAGGNACTGCAGCAGCTGGAATAGAAATATTTCTCTCAACAGCCGT 2579  
QY 481 CCCAGCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATTCAGAGCC 540  
Db 2580 CCCAGCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATTCAGAGCC 2639  
QY 541 TCTCCCTGTCTGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600  
Db 2640 TCTCCCTGTCTGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699  
QY 601 CCTGTAGCTCAGCGTACACTCTGTGAGCAGCAGCGCCGAGCTCGAAGCTCCAGCTGA 660  
Db 2700 CCTGTAGCTCAGCGTACACTCTGTGAGCAGCAGCGCCGAGCTCGAAGCTCCAGCTGA 2759  
QY 661 ACATGTGTCTGCGCCCGGGCGGCCCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 720  
Db 2760 ACATGTGTCTGCGCCCGGGCGGCCCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG 2819  
QY 721 TCCAAATACAGTCTCTTAATGTGTGAACACCTGACATCCCGGAGTTCCAACCGGGTCAAG 780  
Db 2820 TCCAAATACAGTCTCTTAATGTGTGAACACCTGACATCCCGGAGTTCCAACCGGGTCAAG 2879  
QY 781 TGGGAGACATCGCCACTGGSCATCAGCAGCCAGATCCAGCTTCAGCTTGGTCA 840  
Db 2880 TGGGAGACATCGCCACTGGSCATCAGCAGCCAGATCCAGCTTGGTCA 2939  
QY 841 CCAAGACGGGAGCCTGTTGCTACGACATGAGGTGCCAGCTCGGGCATCGACCTGC 900  
Db 2940 CCAAGACGGGAGCCTGTTGCTACGACATGAGGTGCCAGCTCGGGCATCGACCTGC 2999  
QY 901 AGTGCACTATGGCCCTGATGGAGCTTGCCTGGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTGCACTATGGCCCTGATGGAGCTTGCCTGGAGCTGGAGGTCAAGCATGGCCAGC 3059  
QY 961 TGGAGAACAGGCCCTTAACCTGCCCTCCACCGCGGCTCCACACTCCCGGAAAGAGCCCT 1020  
Db 3060 TGGAGAACAGGCCCTTAACCTGCCCTCCACCGCGGCTCCACACTCCCGGAAAGAGCCCT 3119  
QY 1021 TCCTGCTCGGTGACAGTGTGCTCTCAAAACAACAGGCTCAGCCGTTCACAGGGGATYTG 1080  
Db 3120 TCCTGCTCGGTGACAGTGTGCTCTCAAAACAACAGGCTCAGCCGTTCACAGGGGAT-- 3177  
QY 1081 NCCAGCCCCCGGCTCAGAGNTTGGNAACAGGGCCCTCGNACGACNAGGNTGGGG 1140

Db 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGCCCTCG-----CAGCAGCAAGGTGGGG 3230  
QY 1141 CAAGCNAAGATGCTCCAGAGATTTTCAANCTTGAGCCCTGAGCCCTGCCCCANCCCTGCTGAADA 1200  
Db 3231 GCAAGCAAGATGCTCCAGAGATTTTCACTGAGCCCTGCCAC-----CCTGCTGA 3284  
QY 1201 AAACAYTCCGCCACCTGGAAGAGACAGAGAGATGCGNAGAGGATTTNNACCTTGGGGAA 1260  
Db 3285 AAAAAATCCGCCACCTGGAAGAGACAGAGAGAGATGGC---AGGAGTTTACCTGGGGAA 3340  
QY 1261 ACAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTGGCCCTGACCCCTGCTGG 1320  
Db 3341 ACAAAACAGGATCTTTTCTGCCCCCTGCTCCAGT--CGAGTTGGCTGA--CCGCTTG 3396  
QY 1321 ANTCAGTACCATTTGTCAGANCAAGGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGGG 1380  
Db 3397 GATCAGTGACCATTTGTTGGCAGA--CAGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGGG 3455  
QY 1381 GTGGGAGAGCCCTTTCGGCCCTCAACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440  
Db 3456 GTGGGAGAGCCCTTTCGGCCCTCAACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA 3513  
QY 1441 AGGNCCTCAANCTCAGNTTTCAGTGAGAACACAGGTNCCAGCAGTATGCCGCCCGCNTA 1500  
Db 3514 AGGNCCTCAAACTC---AGTTTCAGTGAGAACACAGGT--CAGCAGTATGCCGCCCGTAG 3569  
QY 1501 GGTTAANNNGGGGGCCCTTNAAACCCCTTGCTNGGCCCTNCCAGTNGGCCAGCTCAACCC 1560  
Db 3570 GTTAA---GGGGGCCCTCTAAACCCCTTGCTGSCCTCA--CCTGCCAGCTCA--CCC 3621  
QY 1561 CTTTGGGTGTAGGGGAAAAGATTCCTGACCCCTGGGAAGGCTWCCCTGGTGTAGATACAC 1620  
Db 3622 CTTTGGGTGTAGGGGAAAAGATTCCTGACCCCTGGGAAGGCT--CCTTGGTGTAGATACAC 3680  
QY 1621 CACACTTTTTCAGTGTGTGCAACACAGCTCTGAGTTGACCTCTGTTGTTTCAGCCCAAGGACC 1680  
Db 3681 CACACTTTTTCAGTGTGTGCAACACAGCTCTGAGTTGACCTCTGTTGTTTCAGCCCAAGGACC 3740  
QY 1681 AAAGAAGGTGTGAAGTGAAGTGTCTTCAGTNCCTCCACAGACATGTGCCCTTTGCTGTG 1740  
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTTCAGT--CCCCAGACATGTGCCCTTTGCTGTG 3799  
QY 1741 GCTACCACTCTTCCAGAGCAGCAGGCCCGAGCCCTTCCAGCCCGCAGACTGCCCCAG 1800  
Db 3800 GCTACCACTCTTCCAGAGCAGCAGGCCCGAGCCCTTCCAGCCCGCAGACTGCCCCAG 3859  
QY 1801 ACTCGCTGGCACTCAGTGTCTCTCATCTGTAAGGTGAAGGTGATGCAAGGATATGCTTGA 1860  
Db 3860 ACTCGCTGGCACTCAGTGTCTCTCATCTGTAAGGTGAAGGTGATGCAAGGATATGCTTGA 3919  
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNNAAGNNAAGCAGCAGAGAGACGY 1920  
Db 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNNAAGNNAAGCAGCAGAGAGACG- 3976  
QY 1921 TCCGGGCCCCCAGNCCCCACTNATCAGTGTNCCAGCGTNGGTTNCCAGNAGACACA 1980  
Db 3977 TCCGGGCCCCCAGNCCCCACTNATCAGTGTNCCAGCGTNGGTTNCCAGNAGACAGC- 4032  
QY 1981 GCTNCAAGNATCANCACTGACACTNCACTNGCCCTGCCCCCTNGGCCGANGGAGTACTG 2040  
Db 4033 -----TCAGATCACACTGACACTCACCTGCCCCCTGCCCCCTGG--CCAGAGGGTACTG 4083  
QY 2041 CCGNACGGCACTTTCANCTGTGATGNAACCTCAAGACATTTTCATGGCTNGCCCTTNGG 2100  
Db 4084 CCG-ACGGCACTTTCGACT--CTGATGACCTCAAGACATTTTCATGGCTGCCCCCTCGGCA 4140  
QY 2101 GCAGGNCAGGNCAGGNCAGTGTGACACTGTAGNAGCATANGCAANGCCAGAGATGG 2160  
Db 4141 GGGCAGGNCAGGNCAGTGTGACACT-----GTAGGACATAGCAAGCCAGAGATGG 4190  
QY 2161 GGTGNAAGGNCACAGTCTTTCAGCTGTCCANCACTGATGTGACTTNCCTCAAACTCTTN 2220  
Db 4191 GGTGAA--GGGACACAGTCTTTCAGCTGTCCA--CATGCTGTGACT--CCTCAAACTCTT- 4245



```
QY 1321 ANTCAAGTACCAATTTGTTGGCAGANCAGGGAGACAGCTTCCAGCCCTGGGTGAGAAGGG 1380
Db |||||
3397 GATCAGTGACCAATTTGTTGGCAGA-CAGGGAGACAGCTTCCAGCCCTGGGTGAGAAGGG 3455
QY 1381 GTGGCGAGACCCCTTCGGCCCTTCACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440
Db |||||
3456 GTGGCGAGACCCCTTCGGCCCTTCACCCCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGTA 3513
QY 1441 AGGNNCCCAANCTCAGGNTTTCAGTGCAGAACACAGTNCAGCAGATGCGCCGCCGNTA 1500
Db |||||
3514 AGGGCCCAAACTC---AGGTCAGTGCAGAACACAGGT-CAGCAGATGCGCCGCCCGTAG 3569
QY 1501 GGTTAANNGGGGGCCCTCTNAAACCCCTTGCTNGGCCCTNCCAGCTCAGCTCANCCC 1560
Db |||||
3570 GTTAA---GGGGGCCCTTAACCCCTTGCTGACCTCA---CCTGCCAGCTCA-CCC 3621
QY 1561 CTTTGGGTGTAGGGGAAAGAAATGCCCTGAGACCTGGGAAGGTWCCCTGGGTAGATAAC 1620
Db |||||
3622 CTTTGGGTGTAGGGGAAAGAAATGCCCTGAGACCTGGGAAGGT-CCCTGGGTAGATAAC 3680
QY 1621 CACACTTTTCAGTGTGTTGCACACAGCTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 1680
Db |||||
3681 CACACTTTTCAGTGTGTTGCACACAGCTCCTGAGTTGACCTCTGTTTTCAGCCAGGACC 3740
QY 1681 AAAGAAAGTGTGTAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCTTTTGTGCTG 1740
Db |||||
3741 AAAGAAAGTGTGTAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTTGTGCTG 3799
QY 1741 GGTACACTTTTCCAGAGCAGCAGGCCCCGAGCCCTTCAGGCCCAGCAGCTGCCCCAG 1800
Db |||||
3800 GGTACACTTTTCCAGAGCAGCAGGCCCCGAGCCCTTCAGGCCCAGCAGCTGCCCCAG 3859
QY 1801 ACTCGCTGGCACTAGTTCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTCTGA 1860
Db |||||
3860 ACTCGCTGGCACTAGTTCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTCTGA 3919
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGACGY 1920
Db |||||
3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA---GGAAAGCAGCAGAGAGACG- 3976
QY 1921 TCCGGGCCCCAGNCCCCACCTNATCAGTGTNCCAGCGTGTCTNCCAGNAGACAC 1980
Db |||||
3977 TCCGGGCCCCAGNCCCCACCTATCAGTGT---CCAGCGTGTCTTCCCCAGCAGCAGC- 4032
QY 1981 GCTNCAGNATCANCACTGACACTNCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db |||||
4033 -----TCAGATCACACTGACACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4083
QY 2041 CCGNACGGCAGCTTTGCACTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2100
Db |||||
4084 CG-ACGGCACTTTGCACT-CTGATGATGATGATGATGATGATGATGATGATGATGATG 4140
QY 2101 GGAGGNCAGGNCAGGNCAGTGCATGATGATGATGATGATGATGATGATGATGATGATG 2160
Db |||||
4141 GGCGAGGNCAGGNCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATG 4190
QY 2161 GGTGNAAGGNCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
Db |||||
4191 GGTGAA--GGGACACAGTCTTGAAGTGTGCA-CATGATGATGATGATGATGATGATGATG 4245
QY 2221 NCCAGNATTTCTTAAGATAGCAGTGTGATGATGATGATGATGATGATGATGATGATG 2280
Db |||||
4246 --CCAGATTTCTTAAGATAGCAGTGTGATGATGATGATGATGATGATGATGATGATG 4301
QY 2281 CCCAGGAGCTTATCAGGATCTACAGTGTGATGATGATGATGATGATGATGATGATGATG 2340
Db |||||
4302 CCCAGGAGCTTATCAGGATCTACAGTGTGATGATGATGATGATGATGATGATGATGATG 4359
QY 2341 TGTCTGTAGCTCAACCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db |||||
4360 TGTCTGTAGCTCAACCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4419
```

```
QY 2401 TCCACACTTTGCAAGACAGCCCTTTNGTCCAGCTGTCCACATTAAGTCAAGCTGCTCCC 2460
Db |||||
4420 TCCACACTTTGCAAGACAGCCCTTT-GTCCAGCTGTCCACATTAAGTCAAGCTGCTCCC 4478
QY 2461 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
Db |||||
4479 GGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4538
QY 2521 GGTGTGTAGAGAACTCTTTTGAAGCAATAAAGTGTGGGTGATGACAAATGTTAAAAAA 2578
Db |||||
4539 GGTGTGTAGAGAACTCTTTTGAAGCAATAAAGTGTGGGTGATGACAAATGTTAAAAAA 4596

RESULT 3
US-10-342-887-975
; Sequence 975, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-975

Query Match 70.3%; Score 1850.2; DB 18; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCAGGCCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCAGCCGCTGT 60
Db 2100 CCCCTCTCACAGCCAGGCCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCAGCCGCTGT 2159
QY 61 CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTTGAAGA 120
Db 2160 CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTTGAAGA 2219
QY 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCAATCAAGCCCAATTACC 180
Db 2220 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCCAATTACC 2279
QY 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGGCCAG 240
Db 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGGCCAG 2339
QY 241 CTGAGGAGACAAAGGAGAGAGCCCTTAAGTCCAGCTCCTCTCCACAGAGAGAGAGAGAG 300
Db 2340 CTGAGGAGACAAAGGAGAGAGCCCTTAAGTCCAGCTCCTCTCCACAGAGAGAGAGAGAG 2399
QY 301 AGCCAAACAAGTCTCTCCCTTGAATTTTGAAGAGAGAGAGTCTGGGATGTGGGAACCT 360
Db 2400 AGCCAAACAAGTCTCTCCCTTGAATTTTGAAGAGAGAGAGTCTGGGATGTGGGAACCT 2459
QY 361 TACCTCTGCTCTCCCTGAGAGCCCTTGCAGAAACCCAGCTCACCAGAGCGGAGAG 420
Db |||||
```

Db 2460 TACCTCTGTCTCTCTCCCTGGAGCCAGCCCTGTCAGAAACCCAGCTCACAGAGCGGAAG 2519  
Qy 421 CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCCCTGT 480  
Db 2520 CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCCCTGT 2579  
Qy 481 CCAGGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTGTGTGCTCAGATGACAGCC 540  
Db 2580 CCAGGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTGTGTGCTCAGATGACAGCC 2639  
Qy 541 TCTCCCTGTGGATGACAGTGAAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA 600  
Db 2640 TCTCCCTGTGGATGACAGTGAAGAAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA 2699  
Qy 601 CCCTGAGCTCAGGCGGTACCTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTTCCAGCTGGA 660  
Db 2700 CCCTGAGCTCAGGCGGTACCTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTTCCAGCTGGA 2759  
Qy 661 ACATGGTGTGGCCCGGGGGGCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720  
Db 2760 ACATGGTGTGGCCCGGGGGGCGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2819  
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCCGGAGTTCCACCGGTCAAG 780  
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCCGGAGTTCCACCGGTCAAG 2879  
Qy 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGAGCTTCAGCTTGGTCA 840  
Db 2880 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGAGCTTCAGCTTGGTCA 2939  
Qy 841 CCAAAGACGGGACGCTTGTGCTTACGACATGAGAGTGCAGACTCGGGCATGACCTGC 900  
Db 2940 CCAAAGACGGGACGCTTGTGCTTACGACATGAGAGTGCAGACTCGGGCATGACCTGC 2999  
Qy 901 AGTGCACTGGCCCTGATGGGAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTGCACTGGCCCTGATGGGAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3059  
Qy 961 TGGAGAACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACACTGCGGAAAGACGCT 1020  
Db 3060 TGGAGAACAGGCCCTAACCTCGCCCTCCACCGCGGCTCCACACTGCGGAAAGACGCT 3119  
Qy 1021 TCCTGCTCGGTGACAGATGTCCTGAAACACAGAGCTCAGCGTTCCAGGGGATYTG 1080  
Db 3120 TCCTGCTCGGTGACAGATGTCCTGAAACACAGAGCTCAGCGTTCCAGGGGATYTG 3177  
Qy 1081 NCCAGCCCCCGGCTCAGCAGNTGGAAACAGGCGCTCGNAGCAGNAGTNGGGG 1140  
Db 3178 GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGCGCTCG-----CAGCAGCAAGGTGGG 3230  
Qy 1141 CAAGCNAGAAATGCTCCAGGATTTACANCTGAGCCCTGAGCCCTGCTGAADA 1200  
Db 3231 GCAAGCAAGATGCTCCAGGATTTACACTGAGCCCTGCGCCAC-----CTGCTGA 3284  
Qy 1201 AAAAYTNCGCCACGTGAAGAGACAGAAGAGGATGNCAGAGTTNNACCTYGGGGAA 1260  
Db 3285 AAAAACAATCCGCCAGTGAAGAGACAGAAGAGGATGSC---AGGAGTTACCTGGGGAA 3340  
Qy 1261 ACAAAAAGGATCTTTTCTGCTCCCTGCTCAGTNGAGTTGGCCCTGACCCCTGG 1320  
Db 3341 ACAAAAAGGATCTTTTCTGCTCCCTGCTCAGT--CGAGTTGGCCCTGA--CCCGCTG 3396  
Qy 1321 ANTCAAGTACATTTGTCAGANCAAGGAGAGAGCTTCCAGCCTGGGTCAAGAGG 1380  
Db 3397 GATCAGTGACATTTGTTGGCAGA--CAGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGG 3455  
Qy 1381 GTGGGCGAGCCCTTCGGCCCTCAACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440  
Db 3456 GTGGGCGAGCCCTTCGGCCCTCAACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA 3513  
Qy 1441 AGGNCACCAANTCAGNTTCAAGTGTGAGAAACAGAGTNCAGAGGATGCCCCCAGNTA 1500  
Db 3514 AGGCCCCAAATC---AGGTTTCAAGTGTGAGAAACAGGTT--CAGCAGGATGCCCCCGCTAG 3569

Qy 1501 GGTAAANNNGGGGCCCCCTCTNAAACCCCTTCCCTNGGCTNCACTNGGCGAGCTCANCCC 1560  
Db 3570 GTTAA-----GGGGGCCCTCTAAACCCCTTGCCTGGCCTCA---CCTGGCCAGCTCA-CCC 3621  
Qy 1561 CTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCTWCCCTGCTAGAAATACAC 1620  
Db 3622 CTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCT--CCCTGGTGTAGAAATACAC 3680  
Qy 1621 CACACTTTTTCAGGTTGTTGCAACACAGGCTCCTGAGTTGACCTCTGGTTTCAGCCCAAGACC 1680  
Db 3681 CACACTTTTTCAGGTTGTTGCAACACAGGCTCCTGAGTTGACCTCTGGTTTCAGCCCAAGACC 3740  
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCCCTTTCTGCTG 1740  
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT--CCCCAGACATGTGCCCCCTTTCTGCTG 3799  
Qy 1741 GCTACCACTCTTCCCAAGAGCAGAGCCCGGAGCCCTTTCAGGCCCAAGCTGCCCCAG 1800  
Db 3800 GCTACCACTCTTCCCAAGAGCAGAGCCCGGAGCCCTTTCAGGCCCAAGCTGCCCCAG 3859  
Qy 1801 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGTA 1860  
Db 3860 ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAGGTGAAGGTGATGAGGATATGCTGTA 3919  
Qy 1861 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY 1920  
Db 3920 CAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGCY 3976  
Qy 1921 TCCGGGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGTTTCCCAAGAGCAGCAG- 1980  
Db 3977 TCCGGGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGTTTCCCAAGAGCAGCAG- 4032  
Qy 1981 GCTNAGNCAATCANCACTGACACTNCACTNGCCCTGCCCCCTNGGCCAAGAGGTAAGT 2040  
Db 4033 -----TCAGCATCACTGACACTCACCCTGCCCCCTGCCCCCTG--CCAGAGGGTACTG 4083  
Qy 2041 CCGNAGCGCACTTTTGCACTNCTGATGNACCTCAAAGCACTTTTCATGCTGTCCTCTNNG 2100  
Db 4084 CCG-ACGGCACTTTTGCACT--CTGATGACCTCAAAGCACTTTTCATGCTGTCCTCTGCA 4140  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGGAGCATANGCAAGCCAGAGAGATGG 2160  
Db 4141 GGGCAGGCGCAGGCGAGTGCACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4190  
Qy 2161 GGTGNAAGGANCACAGTCTTGAGCTGCCANCACTGATGACTNCTCTAAACCTCTTN 2220  
Db 4191 GGTGAA--GGGACACAGTCTTGAGCTGTCCA--CATGATGTGACT--CCTCAAACTCTT- 4245  
Qy 2221 NCCAGNATTTCTCTAAGAATAGCANCCCCCTTNCCTTCCCTTGGCCCCAGCTTAGCCTCTTCT 2280  
Db 4246 --CCAGATTTCTCTAAGAATAGCACCCCC--TTCCCCATTTGGCCCCAGCTTAGCCTCTTCT 4301  
Qy 2281 CCAGGGAGCTANTCAGAGCTCACTAGATCAATTAATAGCTGTGNAATCGTCAGGGGG 2340  
Db 4302 CCAGGGAGCTA-CTCAGGACTCAGCTAGCAATTAATAGCTGTG--AATCGTCAGGGGG 4359  
Qy 2341 TGTCTGTAGCTCAACCTCTCGGGCAGGGAGCGCGAGACTCCGTTGGGAGAGAGCTCAT 2400  
Db 4360 TGTCTGTAGCTCAACCTCTCGGGCAGGGAGCGCGAGACTCCGTTGGGAGAGAGCTCAT 4419  
Qy 2401 TCCACATCTTTGCAAGACAGCCTTTTNGTCTCAGCTGTCCACATTTAGTTCAGACTGCTCCC 2460  
Db 4420 TCCACATCTTTGCAAGACAGCCTTTT--GTCCAGCTGTCCACATTTAGTTCAGACTGCTCCC 4478  
Qy 2461 GGGGAGAGAGCCCCCGCCCCCAGCACAATAAGAACTCCAGCCTTGGTACTGACAGTCTG 2520  
Db 4479 GGGGAGAGAGCCCCCGCCCCCAGCACAATAAGAACTCCAGCCTTGGTACTGACAGTCTG 4538  
Qy 2521 GGTGTGTAGAACTCTTTGTAAAGCAATAAAGTTTGGGCTGATGACAAATGTTAAAAA 2578  
Db 4539 GGTGTGTAGAACTCTTTGTAAAGCAATAAAGTTTGGGCTGATGACAAATGTTAAAAA 4596

RESULT 4  
US-10-283-975A-294  
; Sequence 294, Application US/10283975A  
; Publication No. US20040110792A1  
; GENERAL INFORMATION:  
; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
; FILE REFERENCE: CDS 293 PCT  
; CURRENT APPLICATION NUMBER: US/10/283,975A  
; CURRENT FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: 60/340,938  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/338,997  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,081  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/341,012  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-283-975A-294

Query Match 70.3%; Score 1850.2; DB 19; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60  
Db CCCCTCTCACAGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159

QY 61 CTGACGCGAGTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120  
Db CTGACGCGAGTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2219

QY 121 GCCCTTGGGGGAGAAATATAAGAACCAAGACATCCACCGCCAATCAAGCCAATTACC 180  
Db GCCCTTGGGGGAGAAATATAAGAACCAAGACATCCACCGCCAATCAAGCCAATTACC 2279

QY 181 ACCAGACCTTCATGCCAGCGAGAGAGCTTTCCGCAAGGGCCCGAGGCCCGCCAG 240  
Db ACCAGACCTTCATGCCAGCGAGAGAGCTTTCCGCAAGGGCCCGAGGCCCGCCAG 2339

QY 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCAG 300  
Db CTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCAG 2399

QY 301 AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGNACCT 360  
Db AGCCAAACAAGTCTCTCCCTTGACTTTGAGCAAGGAGAGTCTGGGATGTGGNACCT 2459

QY 361 TACCTCTGTCTCCCTGGAGCGAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAG 420  
Db TACCTCTGTCTCCCTGGAGCGAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAG 2519

QY 421 CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAAATATCTCAACAGCCTGT 480  
Db CAACCGTCCCGAGCAGGAACTGACAGAGCTGGAAATAGAAATATCTCAACAGCCTGT 2579

QY 481 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTCTGCTCAGCATCGACGC 540  
Db CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTCTGCTCAGCATCGACGC 2639

QY 541 TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAGGCGCTCTCAAAAGCTCGCGGACA 600  
Db TCTCCCTGTGCGATGACAGTGAAGAAACCCATCAAGGCGCTCTCAAAAGCTCGCGGACA 2699

QY 601 CCCTGAGCTCAGGCGTACACTCTGAGCAGCAGCGGAGGCTGGAAGCTTCCAGCTGGA 660  
Db CCCTGAGCTCAGGCGTACACTCTGAGCAGCAGCGGAGGCTGGAAGCTTCCAGCTGGA 3799

Db 2700 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCCGCGAAGCTCCAGCTGGA 2759  
QY 661 ACATGGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 720  
Db 2760 ACATGGTGTCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAG 2819  
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACATCCCGGAGGTTCACCGGGTCAAG 780  
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACATCCCGGAGGTTCACCGGGTCAAG 2879  
QY 781 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 840  
Db 2880 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 2939  
QY 841 CCAAAGACGGGAGCGCTTTCGCTACGACATGGAAGTGCAGACTCGGGCATTCGACCTGC 900  
Db 2940 CCAAAGACGGGAGCGCTTTCGCTACGACATGGAAGTGCAGACTCGGGCATTCGACCTGC 2999  
QY 901 AGTGCACACTGGCCCTGATGGCAGCTTCGCTGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTGCACACTGGCCCTGATGGCAGCTTCGCTGAGCTGGAGGTCAAGCATGGCCAGC 3059  
QY 961 TGGAGAACAGGCGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 1020  
Db 3060 TGGAGAACAGGCGCCCTAACCCCTGCCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 3119  
QY 1021 TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTG 1080  
Db 3120 TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATTT-- 3177  
QY 1081 NCCAGCCCCCGGCTCARGCAGTGGGAACAGGCGCTCGNCAGCAGNAGNAAAGTTNGGGG 1140  
Db 3178 GCGAGCCCCCGGCTCA--CAGTGGGAACAGGCGCTCG-----CAGCAGCAAGTGGGG 3230  
QY 1141 CAAGNAGATGCTCCAGGATTTACANCTGAGCGCCNTGCCCANCCCTGCTGADA 1200  
Db 3231 GCAAGCAGATGCTCCAGGATTTACACTGAGCCTGCCCCAC-----CCTGCTGA 3284  
QY 1201 AAACATYNNCCGACAGTGAAGAGACAGAGGAGATGNCAGGAGTTNNACCTTGGGGAA 1260  
Db 3285 AAAAACAATCCGCCACGTGAAGAGACAGAGGAGATGNCAGGAGTTNNACCTTGGGGAA 3340  
QY 1261 AAAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTTCGCTGACCTGACCTGG 1320  
Db 3341 AAAAAACAGGATCTTTTCTGCCCCCTGCTCCAGT--CGAGTGGCTCGA--CCGCTTG 3396  
QY 1321 ANTCAGTGACATTTGTCAGANAGGGGAGAGAGCTTCAGAGCTTGGGTCAAGAGG 1380  
Db 3397 GATCAGTGACATTTGTTGGCAGA-CAGGGGAGAGCAGCTTCCAGCTTGGGTCAAGAGG 3455  
QY 1381 GTGGGAGCCCTTCGGCCCTCACCTNCCAGGCTGCTGTGAGAGTGTCAAGTGTGA 1440  
Db 3456 GTGGGAGCCCTTCGGCCCTCACCTNCCAGGCTGCTGTGAGAGTGTCAAGTGTGA 3513  
QY 1441 AGGNCNCAANCTCAGNTTCAGTGCAGAACAGGTCNAGCAGGTATGCCGCCCGNNTA 1500  
Db 3514 AGGCCCAAACTC---AGGTTCACTGACAGAACAGGT-CAGCAGGTATGCCGCCCGCTAG 3569  
QY 1501 GGTAAANNNGGGGCGCTCTNAAACCCCTTGCCTNNGGCTTCCCTGAGCTCAGCTCACC 1560  
Db 3570 GTTAA----GGGGCCCTCTAAACCCCTTGCCTGCGCTCA---CCTGCCAGCTCA--CCC 3621  
QY 1561 CTTTTGGGTGTAGGGGAAAGATGCTGACCTCGCCCTGGGAAGGCTTCCTGCTGAGATACAC 1620  
Db 3622 CTTTTGGGTGTAGGGGAAAGATGCTGACCTCGCCCTGGGAAGGCT--CCCTGGTAGAATACAC 3680  
QY 1621 CACACTTTTCAGGTGTGCAACACAGGCTCTGAGTTGACCTCTGTTTCAGCCCAAGGACC 1680  
Db 3681 CACACTTTTCAGGTGTGCAACACAGGCTCTGAGTTGACCTCTGTTTCAGCCCAAGGACC 3740  
QY 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGTNCCTCAGATGCCAGCATGTGCCCTTTGCTGTG 1740  
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT--CCCCAGACATGTGCCCTTTTCTGCTGTG 3799

```
QY 1741 GCTACCACTCTTCCCGAGAGCAGAGGCCCCGAGGCCCTTCAGGCCACAGCACTGCCCCAG 1800
DB 3800 GCTACCACTCTTCCCGAGAGCAGAGGCCCCGAGGCCCTTCAGGCCACAGCACTGCCCCAG 3859
QY 1801 ACTCGTGGCACTCAGTTCCCTCATCTCTGAAGGTGAAGGGTGATGCAAGGATATGCCTGA 1860
DB 3860 ACTCGTGGCACTCAGTTCCCTCATCTCTGAAGGTGAAGGGTGATGCAAGGATATGCCTGA 3919
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGGNAAGCAGCAGAGAGAGAGCY 1920
DB 3920 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAGCAGCAGAGAGAGAGC- 3976
QY 1921 TCCGGGCCCCCAGNCCCACATNATCAGTGTNCCAGCGTGTCTNGTNGTNNCCCCAGNAGCACA 1980
DB 3977 TCCGGGCCCCCAGNCCCACATNATCAGTGTNCCAGCGTGTCTNGTNGTNNCCCCAGNAGCACA 4032
QY 1981 GCTNAGNCAATCANCACACTGACACTNCACCTTNGCCCTTNGCCCTTNGCCCTTNGCCCTTNG 2040
DB 4033 -----TCGATCATCAGTACACTCACCCTGCCCCCTGCCCCCTGG--CCAGAGGGTACTG 4083
QY 2041 CCGNACGGCACTTTGCACTNTGTATGNACCTCAAGACACTTTTCATGTGCTNGCCCTCTNNG 2100
DB 4084 CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGCACTTTTCATGTGCTNGCCCTCTTGGCA 4140
QY 2101 GCAGGGNAGGGNAGGNCAGTGCACANTCTGTAGNAGCATPANGCAANGCCAGAGATGG 2160
DB 4141 GGGCAGGGCAGGGCAGTGACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190
QY 2161 GGTNAGGGNAGNACAGCTTTGAGCTGTCANCATGATGATGACTNCTCAAACTCTTN 2220
DB 4191 GGTGAA--GGGACACAGCTTTGAGCTGTGCTCA--CATGATGTGACT--CCTCAAACTCTT- 4245
QY 2221 NCCAGNATTTCTTAAGATAAGNACCCCTTNGCCCTTNGCCCTTNGCCCTTNGCCCTTNGCT 2280
DB 4246 --CCAGATTTCTTAAGNATAGCACCCCT--TTCCCAATTTGCCCTGAGCTTNGCTTCTCT 4301
QY 2281 CCCAGGGAGTANTCAGGACTCAGTGTAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340
DB 4302 CCCAGGGAGGCTA--CTCAGGACTCAGCTAGCATTAATCAGCTGTG--AATCGTCAGGGGG 4359
QY 2341 TGTCTGTAGCTCAACTCTCGGGCAGGGAGCGCCGAGACTCGTGGGAGAGCTCAT 2400
DB 4360 TGTCTGTAGCTCAACTCTCGGGCAGGGAGCGCCGAGACTCGTGGGAGAGCTCAT 4419
QY 2401 TCCACATCTTGGCAAGACAGCTTTNGTGTGAGCTGTCCAGATTTGAGTCAAGCTGTCCC 2460
DB 4420 TCCACATCTTGGCAAGACAGCTTTT--GTCAGCTGTCCATTTGAGTCAAGCTGTCCC 4478
QY 2461 GGGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGCGAGCTTTGGTACTGCGAGTCTG 2520
DB 4479 GGGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGCGAGCTTTGGTACTGCGAGTCTG 4538
QY 2521 GGTGTAGAGACTCTTTGTAGCAATAAGTTTGGGGTGATGACAAATGTTAAAAA 2578
DB 4539 GGTGTAGAGACTCTTTGTAGCAATAAGTTTGGGGTGATGACAAATGTTAAAAA 4596
```

## RESULT 5

```
US-10-087-192-887
; Sequence 887, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087.192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
```

```
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 887
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-087-192-887
```

Query Match 67.8%; Score 1783.4; DB 13; Length 4684;

Best Local Similarity 89.3%; Pred. No. 0;

Matches 2297; Conservative 5; Mismatches 183; Indels 87; Gaps 35;

```
QY 1 CCCCTCTCAGAGCCAGCCAGCCATCCAGAGGGCTTGAGGAAAGAGCCCATCCACCGCGTGT 60
DB 2150 CCCCTCTCAGAGCCAGCCAGCCATCCAGAGGGCTTGAGGAAAGAGCCCATCCACCGCGTGT 2209
QY 61 CTCAGCGGAGCTGGGAGGAGGTAACCCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120
DB 2210 CTCAGCGGAGCTGGGAGGAGGTAACCCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2269
QY 121 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCCAATTACC 180
DB 2270 GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCCAATCAAGCCCAATTACC 2329
QY 181 ACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCCCAAGGGGCCCCAGGGCCCCCGGCCAG 240
DB 2330 ACCAGACCTTCCATGCCAGCCGAGAGAGCTTTCCCAAGGGGCCCCAGGGCCCCCGGCCAG 2389
QY 241 CTGAGGAGACACAGGCGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCCCAG 300
DB 2390 CTGAGGAGACACAGGCGAGAGCCCTTAAGCTCCAGCCCTCTCTCCACAGAGCCCCCAG 2449
QY 301 AGCCAAACAAAGTCTCTCCCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 360
DB 2450 AGCCAAACAAAGTCTCTCCCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCT 2509
QY 361 TACCTCTGTCTCTCTCGAGCCAGCCCTCTGCAAGAAACCCAGCTCACCAGAGCGGAAAG 420
DB 2510 TACCTCTGTCTCTCTCGAGCCAGCCCTCTGCAAGAAACCCAGCTCACCAGAGCGGAAAG 2569
QY 421 CAACCGTCCGGAGCAGGAACTGCAGAGCTGGAATATAGAAATATTCTCTCAACAGCTGT 480
DB 2570 CAACCGTCCGGAGCAGGAACTGCAGAGCTGGAATATAGAAATATTCTCTCAACAGCTGT 2629
QY 481 CCAGCCATTTCTCTCGAGGAGCAGAGCAAAATTTCTCTGCTGCTCAGCATCCACAGCC 540
DB 2630 CCAGCCATTTCTCTCGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCCACAGCC 2689
QY 541 TCTCCCTGTCTCGATGACAGTGAGAAAGAACCCATCAAAAGGCCCTCTCAAGGCTCGCGGACA 600
DB 2690 TCTCCCTGTCTCGATGACAGTGAGAAAGAACCCATCAAAAGGCCCTCTCAAGGCTCGCGGACA 2749
QY 601 CCCTGAGCTCAGCGGTACACTCTGAGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTTGA 660
DB 2750 CCCTGAGCTCAGCGGTACACTCTGAGAGCAGCAGCCGAGGCTCGAAGCTCCAGCTTGA 2809
QY 661 ACATGTGTCTGGCCGGGGGGGCCACCGACACCCCAAGCTATTTCATTTGCTGTAAG 720
DB 2810 ACATGTGTCTGGCCGGGGGGGCCACCGACACCCCAAGCTATTTCATTTGCTGTAAG 2869
QY 721 TCCAAATACAGTCTCTTAATGTGAACACCTGACATCCCGGAGTTCCACCGGCTCAAG 780
DB 2870 TCCAAATACAGTCTCTTAATGTGAACACCTGACATCCCGGAGTTCCACCGGCTCAAG 2929
QY 781 TGGGAGACATCGCCATCTGGCATCAGCAGCAGATCCAGCTGACGCTTCACTTGTGTCA 840
DB 2930 TGGGAGACATCGCCATCTGGCATCAGCAGCAGATCCAGCTGACGCTTCACTTGTGTCA 2989
QY 841 CCAAGACGGGCGACCTGTTGCTTACGACATGAGGTTGCCAGACTCGGGCATCCAGCTGC 900
DB 2990 CCAAGACGGGCGACCTGTTGCTTACGACATGAGGTTGCCAGACTCGGGCATCCAGCTGC 3049
```



```
QY 901 AGTGACACTGGCCCTGATGACGCTTCGCTGAGCTGAGGGTCAAGCATGGCCAGC 960
Db 3050 AGTGACACTGGCCCTGATGACGCTTCGCTGAGCTGAGGGTCAAGCATGGCCAGC 3109
QY 961 TGGGAACAGGCGCTTAACCTGCTCCACCGCCGGCTCCACACTGCGGGAAGCAGCCT 1020
Db 3110 TGGGAACAGGCGCTTAACCTGCTCCACCGCCGGCTCCACACTGCGCGG-AAGCAGCCT 3168
QY 1021 TCCTGCTGGTGACAGTGGCTGCTGAAACACAGGCTCAGCGTTCCTCCAGGGATYTG 1080
Db 3169 TCCTGCTGGTGACAGTGGCTGCTGAAACACAGGCTCAGCGTTCCTCCAGGG---ATC 3224
QY 1081 NCCAGCCCCCGGCTCAGCAGTGGGAACACAGGCGCTCGCAGCAGCAAG-----GTGGGG 3277
Db 3225 TGCCAGCCCCCGGCTCAGCAGTGGGACAGGCGCTCGCAGCAGCAAG-----GTGGGG 3277
QY 1141 CAAGCNAGAATGCTCCACAGGATTTTCANCTGAGCCGCTGCCCCAGCCCTGCTGAADA 1200
Db 3278 GCAAGCAGAAATGCTCCACAGGATTTTCACACC---TGAGCCCTGCGCCACCCCTGCTGAGAA 3334
QY 1201 AAACAYTNCGCCACGCTGAGGACAGAGGAGGATGNCAGGAGTTNACCTYGGGAA 1260
Db 3335 AACACT--CCGCCACGTAAGAGACA---GAGGAGGATGGCAGGATTAACCTCGGGAAA 3388
QY 1261 ACAAAACAGGGATCTTNTTCTGCCCCCTGCTCCAGTNGAGTGGCTGACCCCTTGG 1320
Db 3389 CAACAGGATCTTCTGCCCCCTGCTCCAG-----TCGAGTTGGCTGACCCGCTGG- 3440
QY 1321 ANTCAGTGACCAATTTGTTGGCAGANACAGGGGAGAGAGCTTCCAGCCTGGGTGAGAAGG 1380
Db 3441 -ATCAGTGACCAATTTGCTGGCAGA-CAGGGGAGAGCAGCTTCCAGCCTGGGTGAGAAGG 3498
QY 1381 GTGGGAGAGCCTTGGGCCCCCTCACCTNCCAGGCTGCTGTCNAGAGTGTCAAGTGTA 1440
Db 3499 GTGGGAGAGCCTTGGGCCCCCTCACCCCT--CCAGGCTGCTGTG-AGAGTGTCAAGTGTA 3556
QY 1441 AGGNCNCCAAANCTCAGGNTTCAGTGCAGAACAGGTCNAGCAGGATGTCGCCGCCGNTA 1500
Db 3557 AGGCCCAACTC---AGGTTCAAGTGCAGAACAGGT-CAGCAGGATGTCGCCGCCGNTAG 3612
QY 1501 GGTAAANGGGGCCCTCTNAAACCCCTTGGCTNGGCCCTNCACTNGGCCAGCTCANCC 1560
Db 3613 GTTAA---GGGGGCCCTCTAAACCCCTTGGCTNGGCCCTCA---CCTGGCCAGCTCA-CC 3664
QY 1561 CTTTGGGTGAGGGGAAAGATGCTGACCTGGGAAGGCTWCCTCGTGAATATACAC 1620
Db 3665 CTTTGGGTGAGGGGAAAGATGCTGACCTGGGAAGGCT-CCCTGGTGAATATACAC 3723
QY 1621 CACACTTTTCAGGTTGTTGCAACACACAGGTCCTGAGTTGACCTCTGGTTCAGCCAGGACC 1680
Db 3724 CACACTTTTCAGGTTGTTGCAACACACAGGTCCTGAGTTGACCTCTGGTTCAGCCAGGACC 3783
QY 1681 AAGAAGGTGTGAAGTGGTTCTCAGTNCNCCACACATGTCGCCCTTGTGCTG 1740
Db 3784 AAGAAGGTGTGAAGTGGTTCTCAGT-CCCAGACATGTGCCCCCTTGTGCTG 3842
QY 1741 GCTACCACTTTCACAGACAGCAGGCCCGGAGCCCTTTCAGGCCAGCACTGCCCCCAG 1800
Db 3843 GCTACCACTTTCACAGACAGCAGGCCCGGAGCCCTTTCAGGCCAGCACTGCCCCCAG 3902
QY 1801 ACTCGTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGGTGAATCAGGATATGCTGA 1860
Db 3903 ACTCGTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGGTGAATCAGGATATGCTGA 3962
QY 1861 CAGGAACAGTCTGTGATGACATGATCATGCTNAGGNAAGNAAAGCAGCAGAGACGY 1920
Db 3963 CAGGAACAGTCTGTGATGACATGATCATGCTGTA--GGAAAGCAGCAGAGACGCG 4020
QY 1921 TCCGGGCCCCAGNCCCCTNATCAGTGTNCCAGGCTGCTNGTTCNCCAGNAGCACA 1980
Db 4021 TCCGGGCCCCAGNCCCCTNATCAGTGT---CAGGCTGCTGTTCCC-----CAGA 4069
QY 1981 GCTNCAAGNATCANCACTGACATNCAACCCCTNGCCCTGCCCCCTNGGCCANGAGGGTACTG 2040
```

```
Db 4070 GCACAGCTCAGCATCAGCATGACACTCACCTGCGCTGCCCTGG--CCAGAGGGTACTG 4127
QY 2041 CCNNAACGGCACTTTGCACTGACNTCTGATGACCTCAAGCACTTTTCATGCTGCTGCTCTTNG 2100
Db 4128 CCG-ACGGCACTTTGCACT-TCGATG-ACCTCAAGCACTTTTCATGCTGCTGCTCTTNGCA 4184
QY 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAANGCAGGAGATGG 2160
Db 4185 GGGCAGGCGAGGCGAGTGACACTGT-----AGGAGCATAGCAAGCAGGAGATGG 4234
QY 2161 GGTGNAAGGAGNACAGCTTTGAGCTGCTCCANCATGATGATGACNTCTCAAACTCTTN 2220
Db 4235 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4289
QY 2221 NCCAGNATTTCTCTAAGAAATAGCANCCTTNCCTTCCCTGAGCTTAGCCTCTTCT 2280
Db 4290 --CCAGATTTCTCTAAGAAATAGCACCCT--TTCCCTTCCCTGAGCTTAGCCTCTTCT 4345
QY 2281 CCCAGGGAGCTANTCTCAGGACTCACGCTAGCACTAAATCAGCTGTGNAATCGTCAGGGGG 2340
Db 4346 CCCAGGGAGCTA-CTCAGGACTCAGCTAGCACTAAATCAGCTGTG-AATCGTCAGGGGG 4403
QY 2341 TGTCTGCTAGCTCAACCTCTGCGGCGAGGAGCGCCGAGACTCGTGGGAGAGCTCAT 2400
Db 4404 TGTCTGCTAGCTCAACCTCTGCGGCGAGGAGCGCCGAGACTCGTGGGAGAGCTCAT 4463
QY 2401 TCCACATCTTTCGCAAGACAGCTTTNGTCCAGCTGTCCAGTGTCCATTTGAGTCAGACTGCTCCC 2460
Db 4464 TCCACATCTTTCGCAAGACAGCTTTT-GTCCAGCTGTCCATTTGAGTCAGACTGCTCCC 4522
QY 2461 GGGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG 2520
Db 4523 GGGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG 4582
QY 2521 GGTGTAGAGAACTCTTTGTAGCAATAAAGTTTGGGTGATGACAAATGTT 2572
Db 4583 GGTGTAGAGAACTCTTTGTAGCAATAAAGTTTGGGTGATGACAAATGTT 4634
```

## RESULT 6

US-09-871-889-2

; Sequence 2, Application US/09871889

; Patent No. US2002004299A1

; GENERAL INFORMATION:

; APPLICANT: Greene, Warner C.

; APPLICANT: Lin, Xin

; APPLICANT: Gelezuinas, Romas

; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED

; FILE OF INVENTION: BY TNF-ALPHA AND IL-1

; FILE REFERENCE: 30448.61USU1

; CURRENT APPLICATION NUMBER: US/09/871,889

; CURRENT FILING DATE: 2001-06-01

; PRIOR APPLICATION NUMBER: 09/257,703

; PRIOR FILING DATE: 1999-02-25

; PRIOR APPLICATION NUMBER: 60/076,299

; PRIOR FILING DATE: 1998-02-27

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 2844

; TYPE: DNA

; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)

US-09-871-889-2

Query Match 37.1%; Score 977; DB 9; Length 2844;

Best Local Similarity 100.0%; Pred. No. 1.7e-290; Indels 0; Gaps 0;

Matches 977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCCCTCTTCACAGCCCGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 1868 CCCCTCTTCACAGCCCGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 1927
```



```
QY 61 CTGACGGAGCTGGAGGAGGTGAACCGGCACCTACAGCAAGTGGAGGCTCTGAAGA 120
Db 1928 CTGACGGAGCTGGAGGAGGTGAACCGGCACCTACAGCAAGTGGAGGCTCTGAAGA 1987
QY 121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180
Db 1988 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2047
QY 181 ACCAGACCTCCATGCCCCAGCGAGAGAGCTTTCCGCAAGGGGCCCAAGGGCCCCGGCCAG 240
Db 2048 ACCAGACCTCCATGCCCCAGCGAGAGAGCTTTCCGCAAGGGGCCCAAGGGCCCCGGCCAG 2107
QY 241 CTGAGGAGACAAAGGACAGCCCTTAAGCTTCAGCTCCAGCTCTCTCCAGCAGAGCCCCCAG 300
Db 2108 CTGAGGAGACAAAGGACAGCCCTTAAGCTTCAGCTCCAGCTCTCTCCAGCAGAGCCCCCAG 2167
QY 301 AGCCAAACAAGCTCTCTCCCTTGAGCTTTGACCAAGGAGGAGTCTGGGATGTGGCAACCT 360
Db 2168 AGCCAAACAAGCTCTCTCCCTTGAGCTTTGACCAAGGAGGAGTCTGGGATGTGGCAACCT 2227
QY 361 TACCTCTGTCTCTCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCCAGAGCGGAAAG 420
Db 2228 TACCTCTGTCTCTCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCCAGAGCGGAAAG 2287
QY 421 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 480
Db 2288 CAACCGTCCCGAGCAGGAATCGCAGCAGCTGGAAATAGAAATATTCTCAACAGCCTGT 2347
QY 481 CCCAGCATTCTCTCGAGGAGCAGGAGCAAAATCTCTGTGCTCCTCAGCATCGACAGCC 540
Db 2348 CCCAGCATTCTCTCGAGGAGCAGGAGCAAAATCTCTGTGCTCCTCAGCATCGACAGCC 2407
QY 541 TCTCCCTGTTCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 600
Db 2408 TCTCCCTGTTCGATGACAGTGAAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 2467
QY 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGCGAGGCTCGAAGCTCCAGCTGGA 660
Db 2468 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGCGAGGCTCGAAGCTCCAGCTGGA 2527
QY 661 ACATGGTCTGGCCGGGGGGCCACCGACCCCAAGCTATTCTCAATGGTGAAG 720
Db 2528 ACATGGTCTGGCCGGGGGGCCACCGACCCCAAGCTATTCTCAATGGTGAAG 2587
QY 721 TCCAAATACAGTCTCTTAATGGTGAAACCTTGACATCCCGGAGTTCACCGGTCAAAG 780
Db 2588 TCCAAATACAGTCTCTTAATGGTGAAACCTTGACATCCCGGAGTTCACCGGTCAAAG 2647
QY 781 TGGAGACATCGCACTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTCA 840
Db 2648 TGGAGACATCGCACTGGCATCAGCAGCCAGATCCCGAGCTGCAGCTTCAGCTTGGTCA 2707
QY 841 CCAAGACGGCAGCCCTGTTGCTTACACATGGAGTGCAGACTCGGGCATCGACCTGC 900
Db 2708 CCAAGACGGCAGCCCTGTTGCTTACACATGGAGTGCAGACTCGGGCATCGACCTGC 2767
QY 901 AGTGCACTTGGCCCTCGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960
Db 2768 AGTGCACTTGGCCCTCGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 2827
QY 961 TGGAGAACAGCCCTAA 977
Db 2828 TGGAGAACAGCCCTAA 2844
```

RESULT 7

```
US-10-087-192-886
; Sequence 886, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
```

```
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 886
; LENGTH: 73967
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-087-192-886
```

```
Query Match 37.08; Score 973.4; DB 13; Length 73967;
Best Local Similarity 84.4%; Pred. No. 7.6e-289;
Matches 1487; Conservative 5; Mismatches 183; Indels 87; Gaps 35;
```

```
QY 811 AGATCCCGAGCTGCAGCCTTCAGCTTGTTCACCAAGACGGGACGCTTTCGCTACGACA 870
Db 62243 AGATCCCGAGCTGCAGCCTTCAGCTTGTTCACCAAGACGGGACGCTTTCGCTACGACA 62302
QY 871 TGGAGGTGCCAGACTCTGGGCATCGACCTGCAGTGCACTTGGCCCCCTGATGGCAGTTTCG 930
Db 62303 TGGAGGTGCCAGACTCTGGGCATCGACCTGCAGTGCACTTGGCCCCCTGATGGCAGTTTCG 62362
QY 931 CTTGGAGCTGGAGGTTCAAGCATGGCCAGCTGGGAGAACAGGCCCTTAACCTGCCCTCCAC 990
Db 62363 CTTGGAGCTGGAGGTTCAAGCATGGCCAGCTGGGAGAACAGGCCCTTAACCTGCCCTCCAC 62422
QY 991 CGCGGGCTCCACACTGCGGAAAGACGCTTCTGCTCGGTGCACGATGCTGCCCTGAAA 1050
Db 62423 CGCGGGCTCCACACTGCGCGG--AAGCAGCTTCTGCTCGGTGCACGATGCTGCCCTG-AA 62480
QY 1051 ACACAGCTCAGCGTTTCCAGGGGATYTNCCACGCCCCCGGCTCARGAGTGGGAACC 1110
Db 62481 ACACAGCTCAGCGTTTCCAGGGG---ATCTGCCAGCCCCCGGCTCAGCAGTGGGACCAG 62537
QY 1111 AGGCGCTCGNACGNAGCNAAAGTNGGGGGCAAGCNAGATGCTTCCAGGATTTACAN 1170
Db 62538 GGCTTCGACAGCAAG--GTGGGGGCAAGCAGAAATGCTTCCAGGATTTACAC 62590
QY 1171 CCTGAGCCTNTGCCANCCCTGCTGAADAAAACAYTNCCGCCACGTCGAGAGACAGAG 1230
Db 62591 C---TGAGCCCTGCCACCCTGCTGAGAAAACACT--CCGCCACGTCGAGAGACA---- 62641
QY 1231 GAGGATGGNCAGGAGTTNNACCTYGGGGAACAAAACAGGGATCTTTTCTGCCCCCTGC 1290
Db 62642 GAGGAGATGGCAGGAGTTTACCTCGGGAACAAACAGGATCTTCTGCTGCTGCTCCAG- 62700
QY 1291 TCAGTNCAGATTGGCTGNACCCGCTTGGANTCAGTGACATTTGTTGGCAGANACAGG 1350
Db 62701 -----TCGAGTTGGCTGCA--CCCGCTTGGATCAGTGACCATTTGCTGGGCGA-CAGGG 62751
QY 1351 GAGACAGCTTCCAGCTGGGTGAGAGGGGTGGGAGCCCTTCGCCCCCTCACCCCTNC 1410
Db 62752 GAGACAGCTTCCAGCTGGGTGAGAGGGGTGGGAGCCCTTCGCCCCCTCACCCCT-C 62810
QY 1411 CAGGCTGCTGTGNAGAGTGTCAAGTGTGTAAGGNGCCAAANCTCAGGNTTCAGTGCGAGA 1470
Db 62811 CAGGCTGCTGTG-AGAGTGTCAAGTGTGTAAGGGGCCAAATC---AGGTTTCAGTGCGAGA 62866
QY 1471 ACCAGTNCAGCAGGATGTCGCCCGCCGNTAGGTTAAANNNGGGGCCCTCTNAAACCCCTTG 1530
Db 62867 ACCAGGT-CAGCAGGATGTCGCCCGCCGNTAGGTTAA-----GGGGGCCCTCTTAAACCCCTTG 62921
QY 1531 CTTNGGCCCTNCACCTTNGGCCAGCTCANCCCTTTTGGGTGTAGGGGAAAGATGCTGA 1590
Db 62922 CTTGGCCCTCA---CCTGGCCAGCTCA-CCTCTTTTGGGTGTAGGGGAAAGATGCTGA 62977
QY 1591 CCCTGGGAAGGCTWCCTCTGTGTAAGATACACCACTTTTTCAGGTTTGTGCAACACAGGTC 1650
```



Db	758	ACGCTCCGGGCCCCAGCCCCACTATCAGTGT---CGAGGTGCTGTTGCC-----	806
Qy	1977	CACAGCTNCAGNCATCANCACTGACACTNCACCTTCGCTGCCCCCTTNGCCANGAGGT	2036
Db	807	CAGAGCACAGCTCAGCATCACTGACACTCACCTGCTCCCTGCCCCCTGG--CGAGGGT	864
Qy	2037	ACTGCCGNCAGGCACTTTGCACTCTGATGNACTCAAGACACTTTTCATGGCTNGCCCTC	2096
Db	865	ACTGCCG-ACGGCACTTTGCACT--CTGATGACCTCAAGCACTTTTCATGGCTGCCCTCT	921
Qy	2097	TNNGCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAAGCCAGGAG	2156
Db	922	GGCAGGCGAGGGCAGGGCAGTGACACT-----GTAGGAGCATAGCAAGCCAGGAG	971
Qy	2157	ATGGGGTGNAAGGGANCAACAGTCTTGAGCTGTCCANCATGCATGTGACTNCCCTCAAACCT	2216
Db	972	ATGGGGTGAA--GGGACACAGTCTTGAGCTGTGCA-CATGCATGTGACT--CTCAAACT	1027
Qy	2217	CTTNCCAGNATTTCTCTAAGAAATAGCANCCCCCTTNCCTTCCCTGAGCTTGAAGCTC	2276
Db	1028	CTT---CCAGATTTCTCTAAGAAATAGCACCCCC--TTCCCCATTTGCCCCAGCTTAGCCTC	1082
Qy	2277	TTCTCCAGGGGAGCTANTCTCAGGACTCAGCTAGTGCATTAATCAGCTGTGNAATCTGCAG	2336
Db	1083	TTCTCCAGGGGAGCTA-CTCAGGACTCAGCTAGCATTAATCAGCTGTG-AACTCGCTCAG	1140
Qy	2337	GGGGTGTCGTCTAGCCTCAACCTCTCGGGCAGGGGACGCCGAGACTCCCGTGGGAGAAC	2396
Db	1141	GGGGTGTCGTCTAGCCTCAACCTCTCGGGCAGGGGACGCCGAGACTCCCGTGGGAGAAC	1200
Qy	2397	TCATTTCCACATCTTTGCCAAGACAGCCTTTNGTCCAGCTGTCCACATTTGAGTCTGC	2456
Db	1201	TCATTTCCACATCTTTGCCAAGACAGCCTTT-GTCCAGCTGTCCACATTTGAGTCTGC	1259
Qy	2457	TCCGGGGAGAGAGCCCCGGCCCCCAGCATAAAGAACTTGCAGCCTTGGTACTGCAGAG	2516
Db	1260	TCCGGGGAGAGAGCCCCGGCCCCCAGCATAAAGAACTTGCAGCCTTGGTACTGCAGAG	1319
Qy	2517	CTCTGGGTTGTAGAGAACTCTTTGTAAAGCAATAAGTTTGGGGTGATGACAAATGTTAAAA	2576
Db	1320	CTCTGGGTTGTAGAGAACTCTTTGTAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAA	1379
Qy	2577	AAA 2579	
Db	1380	AAA 1382	

```

RESULT 9
US-10-087-192-884
; Sequence 884, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: MORRIS, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 884
; LENGTH: 2828
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-884

```

Query Match 25.6%; Score 672.8; DB 13; Length 2828;

Best Local Similarity		81.8%	Pred. No. 1.6e-196;
Matches	798;	Conservative	0; Mismatches 157; Indels 21; Gaps 1;
QY	1	CCCCCTCTCAGAGCCCGAGGCCATCCAAAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	60
DB	1873	CACCCCTCAGAGCCCGAGCCATCCAGAGGGGCTGAGAAAGAGCCCGTCCACCGAGCAT	1932
QY	61	CTGAGCGGAGCTGGGAGGGAAGTGAACCGGGCAGCTACAGCAAGTGGGAGGTCTGAAGA	120
DB	1933	CTGCATGGAGCTTCGGAGGAAAGTGCGCAAGGCACCTACAGGAAGTGGGAGGTCTGAAA	1992
QY	121	GCCCTTGGAGGGAGAAATATAAGAACCAAGACATCCACGCCCAAAATCAAGCCCAATTACC	180
DB	1993	GCCCTTGGAAAGGAGAAATATAAGAACCAAGACCTCCACCCCAAGACCAAGCCACCTGCC	2052
QY	181	ACCAGACCTCTCATGCCAGCGGAGAGAGCTTTTCGCCAAGGGGCCAGGGCCCCGGCCAG	240
DB	2053	ACCAGACCTTACTACTCCGCGAGAGA-----GNAACCCACAG	2091
QY	241	CTGAGGAGACAAAGGAGAGCCCTTAAGTTCAGAGCTCTCTCTCCACAGAGAGCCCCAG	300
DB	2092	CCAAAGGCCAACAGAGCGGGGCTCTCAGAGCTCAGCCCTCTCTACCGCCAGAACACAG	2151
QY	301	AGCCAAACAAGTCTCCCTCCCTTGACTTTGACGAGGAGGAGTCTGGATCTGGGAACCCCT	360
DB	2152	AACCGAACAAGCGCCAGCCCTGAACCTGAGCAAGGAGGAGTCTGGCACAATGGGAAACCC	2211
QY	361	TACCTCTGTCTCCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420
DB	2212	TGCCTCTGTCTCCTCTGGAGCCAGCCACTGCCAAAGGCCCGAGCTTCCAGACCGGAGG	2271
QY	421	CAACCGTCCGGAGCAGGAACGTGAGCAGCTGGAAATAGAAATTATTCCTCAACAGCCTGT	480
DB	2272	CAACCTTTGCCAGAGCTGGAGTACAGCAACTGGAGATAGAACTGTTTCTCAACAGACCTGT	2331
QY	481	CCAGCCATTTCTCTGGAGGAGCAGAGACAAATCTCTCGTCGCTCAGATCGACAGCC	540
DB	2332	CCAGCCGTTCTCTCTGGAGGAACAGAAACAAATCTCTCTCGCTCAGATCGACAGCC	2391
QY	541	TCTCCCTCTCGGATGACAGTGAGAAGAACCCATCAAAAGGCTCTCTCAAGCTCTCGGGACA	600
DB	2392	TCTCGCTCTCAGATGACAGTGAAAGAAATCCATCGAAGGCTCTCTCAGAGCTCAGGGACA	2451
QY	601	CCCTGAGCTCAGGGTACATCTCTGGAGCAGCAGCGCGGAGGTCGAAGCTCCAGCTGGA	660
DB	2452	CCCTGAGTTCTGGCGTCTCACTCTTTGGAACAGCCAAAGCTGAGGCAAGAAACCTGCAAGCTGCA	2511
QY	661	ACATGTGTCTGGCCGGGGGGCCCAACCAACCAACCCCAAGCTATTTCAATGGTCTGAAG	720
DB	2512	GCAGCGGCTTGGCCGGGGGGCGGCTACTGACATCCCGAGCTACTTCAACGGGGTCAAGG	2571
QY	721	TCCAAATACAGTCTCTTAATGGTGAAACACTGACACATCCGGGAGTTCACCGGGTCAAAG	780
DB	2572	TCCAGATCCAGTCTCTCAATGGCGAACACTGATATCCGGGAATTCACCGCGTCAAAG	2631
QY	781	TGGGAGACATCGGCATCTGGGCATCAGCAGCCAGATCCAGCTGCAAGCTTTCAGCTTGTGCA	840
DB	2632	TGGGAGACATTCGCCCGGCATCAGCAGCCAGATCCCAAGCCACAGCTTTTCAGCTTGTGTA	2691
QY	841	CCAAAGACGGGACGCTGTTTCGCTACGACATGGAGGTCAGACTCGGGGATCGACCTGC	900
DB	2692	CCAAAGATGACAGCCTGTTTGTCTATGACATGGAGGTGCCAGACTCGGGGATCGACCTGC	2751
QY	901	AGTGCAACATCGGCCCTTGATGGCAGCTTTCGCTCTGGAGCTGGAGGTCGAAGCTCGGCCAGC	960
DB	2752	AGTGCAACCTCGGCCCTTGATGGCAGCTTTCGCTCTGGAGCTGGAGGTCGAAGCTCGGCCAGC	2811
QY	961	TGGAGAACAGGCCCTTA	976
DB	2812	TGGAGAACCGACCTTA	2827

RESULT 10

```
US-10-108-260A-1754/c
; Sequence 1754, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1754
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1754
Query Match 13.3%; Score 349.4; DB 17; Length 2475;
Best Local Similarity 99.7%; Pred. No. 1.2e-96;
Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 102 CAAAGTGGGAGGCTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCG 161
DB 2179 CTAGTGGGAGGCTCTGAGAGCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCG 2120
QY 162 CCAATCAAGCCAAATTACACAGACCTTCATGCCCCAGCCGAGAGAGCTTTCGCCAAGG 221
DB 2119 CCAATCAAGCCAAATTACACAGACCTTCATGCCCCAGCCGAGAGAGCTTTCGCCAAGG 2060
QY 222 GCCCAGGCCCCGCCAGCTCAGAGACCAAGCAGAGCCCTTAAGCTCCAGCCTCCT 281
DB 2059 GCCCAGGCCCCGCCAGCTCAGAGACCAAGCAGAGCCCTTAAGCTCCAGCCTCCT 2000
QY 282 CTCACACAGAGCCCCCAGAGCAACAAAGTCTCTCTCTTCTGACCTTTGAGCAAGAGGAG 341
DB 1999 CTCACACAGAGCCCCCAGAGCAACAAAGTCTCTCTCTTCTGACCTTTGAGCAAGAGGAG 1940
QY 342 TCTGGATGTGGAAACCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCC 401
DB 1939 TCTGGATGTGGAAACCTTACTCTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCC 1880
QY 402 AGCTACACAGAGCGGAAAGCAACCGTCCCGAGCAGGAACTGCAGCAGCTG 452
DB 1879 AGCTACACAGAGCGGAAAGCAACCGTCCCGAGCAGGAACTGCAGCAGCTG 1829
RESULT 11
US-09-918-995-2585
; Sequence 2585, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2585
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(499)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2585
Query Match 9.2%; Score 242.2; DB 10; Length 499;
Best Local Similarity 90.9%; Pred. No. 8.3e-64;
Matches 288; Conservative 1; Mismatches 23; Indels 5; Gaps 3;
QY 811 AGATCCAGCTGCAGCCTTCAGCTTGGTCAACCAAGACGGGAGCAGCTGTCGTACGACA 870
DB 182 AGATCCAGCTGCAGCCTTCAGCTTGGTCAACCAAGACGGGAGCAGCTGTCGTACGACA 241
QY 871 TGGAGGTGCAGACTCGGGCATCTGCAGCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTCG 930
DB 242 TGGAGGTGCAGACTCGGGCATCTGCAGCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTCG 301
QY 931 CCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCAC 990
DB 302 CCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCAC 361
QY 991 CGCCGGCTCCACACTCCCGAAGAGCAGCTTCTCTGCTCGGTGCACGATGCTGCTCCCTGAAA 1050
DB 362 CGCCGGCTCCACACTCGCCGG--AAGCAGCCTTCTCTGCTCGGTGCACGATGCTGCTCCCTG-AA 419
QY 1051 ACACAGGCTCAGCGCTTCCAGGGGATYTGNCAGCCCCCGGCTCARGAGNTGGGAACC 1110
DB 420 ACACAGGCTCAGCGCTTCCAGGGGATYTGNCAGCCCCCGGCTCARGAGNTGGGAACCAG 476
QY 1111 AGGCGCTCGNCAGNAG 1127
DB 477 GGCCTCGCAGCAGCAAG 493
RESULT 12
US-09-783-590-11852
; Sequence 11852, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11852
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (7)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (28)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (61)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (95)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (100)
; OTHER INFORMATION: n equals a,t,g, or c
```

; NAME/KEY: misc feature  
; LOCATION: (130)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (195)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (228)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (247)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (268)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (277)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (287)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc feature  
; LOCATION: (293)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-783-590-11852

Query Match 7.3%; Score 193; DB 9; Length 296;  
Best Local Similarity 88.2%; Pred. No. 1.1e-48;  
Matches 224; Conservative 0; Mismatches 28; Indels 2; Gaps 2;  
  
Qy 1558 CCCCTTTGGGTAGGGGAAAGAAATGCCCTGACCCCTGGGAAGGCTWCCTGGTAGAATA 1617  
Db 43 CCCCTTTGGGTAGGGGAAAGAAAGAAATGCCCTGACCCCTGGGAAGGCTWCCTGGTAGAATA 102  
  
Qy 1618 CACCACACTTTTTCAGGTGTTGCAACACAGAGTCTCTGAGTTCACCTCTGGTTTCAGCCCAAGG 1677  
Db 103 CACCACACTTTTTCAGGTGTTGCAACANAGTCTCTGAGTTCACCTCTGGTTTCAGCCCAAGG 162  
  
Qy 1678 ACCAAGAGTGTGTAAGTGAAGTGTTCAGTNCCTGAGATGCCATGCCCCCTTTG-CT 1736  
Db 163 ACCAAGAGTGTGTAAGTGAAGTGTTCAGTNCCTGAGATGCCATGCCCCCTTTGCT 222  
  
Qy 1737 GCTGGTACCCTCTTCCACAGAGCAGAGCCCGAGCCCC-TTCAGGCCACGACTGC 1795  
Db 223 GCTGGTACCCTCTTCCACAGAGCAGAGCCCGAGCCCC-TTCAGGCCACGACTGC 282  
  
Qy 1796 CCCAGACTCGCTGG 1809  
Db 283 CCCANATTCGNTGG 296

RESULT 13  
US-10-087-192-883  
; Sequence 883, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 52945200122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 883

; LENGTH: 42566  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)\_(42566)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-883

Query Match 7.2%; Score 188.8; DB 13; Length 42566;  
Best Local Similarity 73.9%; Pred. No. 1.4e-46;  
Matches 263; Conservative 0; Mismatches 72; Indels 21; Gaps 1;  
  
Qy 104 AGTGGAGGCTCTGAAGAGCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACGCC 163  
Db 27368 AGTGGAGGCTCTGAAGAGCCCTTGGAAAGCCCTTGGAAAGGAGATATAAAGAACCAAGACATCCACGCCA 27427  
  
Qy 164 AATCAAGCCCAATTACCAACAGACCCCTCCATGCCCGCCGAGAGAGCTTTCCGCCAAGGCC 223  
Db 27428 AGACCAAGCCCACTGCCACAGACCCCTTACTCTCCGCCGAGAGA----- 27472  
  
Qy 224 CCCAGGCCCGCCGCGAGCTGAGGAGACAAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTC 283  
Db 27473 -----GAACCCACAGCCCAAGGCCCAACAGAGCGGGCTCTGAGCCTCAGCCTCTCT 27526  
  
Qy 284 CCCACAGAGCCCCCAGAGCCAAACAAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTC 343  
Db 27527 ACCGCCAGAACCAACCAAGCCGAGCAAGCCGAGCAAGCCCTGAAACCTGAGCAAGGAGGAGTC 27586  
  
Qy 344 TGGATGTGGAAACCTTACT 403  
Db 27587 TGGCACAATGGAAACCCCTTCT 27646  
  
Qy 404 CTACACAGAGCGGAAACCAACCCGTCGCCGAGAGCAAACTGCAGCAGCTGGAAATAG 459  
Db 27647 CTTCACAGAGCGGAGGCAACCTTGCACAGAGCTGGAGCTGAGCACTGAGAGATAG 27702

RESULT 14  
US-10-357-930-22052/c  
; Sequence 22052, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 62232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22052  
; LENGTH: 455  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-357-930-22052

Mon Jun 13 13:27:12 2005

Query Match 3.4%; Score 88.4; DB 20; Length 455;  
 Best Local Similarity 98.9%; Pred. No. 2.6e-16;  
 Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 458 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 517  
 Db |||||||  
 227 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 168  
 QY 518 CTCGTGCTCAGCATCGACGCTCTCCCT 547  
 Db |||||||  
 167 CTCGTGCTCAGCATCGACGCTCTCCCT 138

RESULT 15  
 US-10-357-930-27909/c  
 ; Sequence 27909, Application US/10357930  
 ; Publication No. US20040259086A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schlegel, Robert  
 ; APPLICANT: Endege, Wilson  
 ; APPLICANT: Monahan, John  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
 ; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
 ; FILE REFERENCE: MRI-007BCN  
 ; CURRENT APPLICATION NUMBER: US/10/357,930  
 ; CURRENT FILING DATE: 2003-02-04  
 ; PRIOR APPLICATION NUMBER: 09/785,276  
 ; PRIOR FILING DATE: 2003-02-16  
 ; PRIOR APPLICATION NUMBER: 60/183,319  
 ; PRIOR FILING DATE: 2000-02-17  
 ; PRIOR APPLICATION NUMBER: 60/189,862  
 ; PRIOR FILING DATE: 2000-03-16  
 ; PRIOR APPLICATION NUMBER: 60/207,454  
 ; PRIOR FILING DATE: 2000-05-25  
 ; PRIOR APPLICATION NUMBER: 60/211,314  
 ; PRIOR FILING DATE: 2000-06-09  
 ; PRIOR APPLICATION NUMBER: 60/219,007  
 ; PRIOR FILING DATE: 2000-07-18  
 ; PRIOR APPLICATION NUMBER: 60/255,281  
 ; PRIOR FILING DATE: 2000-12-13  
 ; NUMBER OF SEQ ID NOS: 62232  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 27909  
 ; LENGTH: 455  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-357-930-27909

Query Match 3.4%; Score 88.4; DB 20; Length 455;  
 Best Local Similarity 98.9%; Pred. No. 2.6e-16;  
 Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 458 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 517  
 Db |||||||  
 227 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGGAGGAGCAAAATTCCT 168  
 QY 518 CTCGTGCTCAGCATCGACGCTCTCCCT 547  
 Db |||||||  
 167 CTCGTGCTCAGCATCGACGCTCTCCCT 138

Search completed: June 11, 2005, 06:13:06  
 Job time : 1014.86 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 6580.54 Seconds  
(without alignments)  
16321.674 Million cell updates/sec

Title: US-09-155-676B-3  
Perfect score: 2631  
Sequence: 1 ccccttcacagccagccg.....accgtcgacctcgagggggg 2631

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1:	/cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq:*
2:	/cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq:*
3:	/cgn2_6/ptodata/1/pna/PCTUS3_COMB.seq:*
4:	/cgn2_6/ptodata/1/pna/US06_COMB.seq:*
5:	/cgn2_6/ptodata/1/pna/US07_COMB.seq:*
6:	/cgn2_6/ptodata/1/pna/US080_COMB.seq:*
7:	/cgn2_6/ptodata/1/pna/US081_COMB.seq:*
8:	/cgn2_6/ptodata/1/pna/US082_COMB.seq:*
9:	/cgn2_6/ptodata/1/pna/US083_COMB.seq:*
10:	/cgn2_6/ptodata/1/pna/US084_COMB.seq:*
11:	/cgn2_6/ptodata/1/pna/US085_COMB.seq:*
12:	/cgn2_6/ptodata/1/pna/US086_COMB.seq:*
13:	/cgn2_6/ptodata/1/pna/US087_COMB.seq:*
14:	/cgn2_6/ptodata/1/pna/US088_COMB.seq:*
15:	/cgn2_6/ptodata/1/pna/US089_COMB.seq:*
16:	/cgn2_6/ptodata/1/pna/US090_COMB.seq:*
17:	/cgn2_6/ptodata/1/pna/US091_COMB.seq:*
18:	/cgn2_6/ptodata/1/pna/US092A_COMB.seq:*
19:	/cgn2_6/ptodata/1/pna/US092B_COMB.seq:*
20:	/cgn2_6/ptodata/1/pna/US093A_COMB.seq:*
21:	/cgn2_6/ptodata/1/pna/US093B_COMB.seq:*
22:	/cgn2_6/ptodata/1/pna/US094_COMB.seq:*
23:	/cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
24:	/cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
25:	/cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
26:	/cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
27:	/cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
28:	/cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
29:	/cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
30:	/cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
31:	/cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
32:	/cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
33:	/cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
34:	/cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
35:	/cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
36:	/cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
37:	/cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
38:	/cgn2_6/ptodata/1/pna/US098D_COMB.seq:*
39:	/cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
40:	/cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
41:	/cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
42:	/cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
43:	/cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44:	/cgn2_6/ptodata/1/pna/US099F_COMB.seq:*
45:	/cgn2_6/ptodata/1/pna/US099G_COMB.seq:*
46:	/cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
47:	/cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
48:	/cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
49:	/cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
50:	/cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
51:	/cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
52:	/cgn2_6/ptodata/1/pna/US103A_COMB.seq:*
53:	/cgn2_6/ptodata/1/pna/US103B_COMB.seq:*
54:	/cgn2_6/ptodata/1/pna/US104A_COMB.seq:*
55:	/cgn2_6/ptodata/1/pna/US104B_COMB.seq:*
56:	/cgn2_6/ptodata/1/pna/US105A_COMB.seq:*
57:	/cgn2_6/ptodata/1/pna/US105B_COMB.seq:*
58:	/cgn2_6/ptodata/1/pna/US106A_COMB.seq:*
59:	/cgn2_6/ptodata/1/pna/US107A_COMB.seq:*
60:	/cgn2_6/ptodata/1/pna/US107B_COMB.seq:*
61:	/cgn2_6/ptodata/1/pna/US107C_COMB.seq:*
62:	/cgn2_6/ptodata/1/pna/US107D_COMB.seq:*
63:	/cgn2_6/ptodata/1/pna/US108A_COMB.seq:*
64:	/cgn2_6/ptodata/1/pna/US108B_COMB.seq:*
65:	/cgn2_6/ptodata/1/pna/US109A_COMB.seq:*
66:	/cgn2_6/ptodata/1/pna/US109B_COMB.seq:*
67:	/cgn2_6/ptodata/1/pna/US109C_COMB.seq:*
68:	/cgn2_6/ptodata/1/pna/US110_COMB.seq:*
69:	/cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
70:	/cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
71:	/cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
72:	/cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
73:	/cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
74:	/cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
75:	/cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
76:	/cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
77:	/cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
78:	/cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
79:	/cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
80:	/cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
81:	/cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
82:	/cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
83:	/cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
84:	/cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
85:	/cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
86:	/cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
87:	/cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
88:	/cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
89:	/cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
90:	/cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
91:	/cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
92:	/cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*
93:	/cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*
94:	/cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
95:	/cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
96:	/cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
97:	/cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
98:	/cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
99:	/cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
100:	/cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
101:	/cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
102:	/cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
103:	/cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
104:	/cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
105:	/cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
106:	/cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
107:	/cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
108:	/cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
109:	/cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
110:	/cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
111:	/cgn2_6/ptodata/1/pna/US6041_COMB.seq:*
112:	/cgn2_6/ptodata/1/pna/US6042_COMB.seq:*
113:	/cgn2_6/ptodata/1/pna/US6043_COMB.seq:*
114:	/cgn2_6/ptodata/1/pna/US6044_COMB.seq:*
115:	/cgn2_6/ptodata/1/pna/US6045_COMB.seq:*
116:	/cgn2_6/ptodata/1/pna/US6046_COMB.seq:*





QY 241 CTGAGGAGACAACAGGAGAGAGCCCTAAGCTCCAGCTCCTCTCTCCACAGAGAGCCCCAG 300  
Db 241 CTGAGGAGACAACAGGAGAGAGCCCTAAGCTCCAGCTCCTCTCTCCACAGAGAGCCCCAG 300  
QY 301 AGCCAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATGTGGAAACCT 360  
Db 301 AGCCAAAACAAGTCTCCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATGTGGAAACCT 360  
QY 361 TACTCTGTCTCTCTGGAGCAGCCTCTGGAGCAGCCTCTGGAGAAACCCAGCTCACCAGAGCGGAAAG 420  
Db 361 TACTCTGTCTCTCTGGAGCAGCCTCTGGAGCAGCCTCTGGAGAAACCCAGCTCACCAGAGCGGAAAG 420  
QY 421 CAACGGTCCCGAGCAGAGAACTGACAGAGCTGGAAATAGAAATTAATCTCAACAGCCTGT 480  
Db 421 CAACGGTCCCGAGCAGAGAACTGACAGAGCTGGAAATAGAAATTAATCTCAACAGCCTGT 480  
QY 481 CCCAGCAATTTCTCTGGAGAGCAGGAGCAAAATTTCTCTCGTGTCTCAGCATCAGCAGCC 540  
Db 481 CCCAGCAATTTCTCTGGAGAGCAGGAGCAAAATTTCTCTCGTGTCTCAGCATCAGCAGCC 540  
QY 541 TCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 600  
Db 541 TCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAAAGGCTCTCAAAAGCTTCGGGGACA 600  
QY 601 CCCTGAGCTCAGGCGTACATCTCTGGAGAGCAGCAGCGGAGGCTCGAAGCTCCAGCTGGA 660  
Db 601 CCCTGAGCTCAGGCGTACATCTCTGGAGAGCAGCAGCGGAGGCTCGAAGCTCCAGCTGGA 660  
QY 661 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG 720  
Db 661 ACATGGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAG 720  
QY 721 TCCAAATACAGTCTCTTAATGGTGAACAACCTGACATCCGGGAGTTCACCGGTCAAAG 780  
Db 721 TCCAAATACAGTCTCTTAATGGTGAACAACCTGACATCCGGGAGTTCACCGGTCAAAG 780  
QY 781 TGGAGACATCGCACTGGGATCAGCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGTGCA 840  
Db 781 TGGAGACATCGCACTGGGATCAGCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGTGCA 840  
QY 841 CCAAGACGGCAGCCTTTCGCTACGACATGGAGGTCCAGACTCGGGCATCGACCTGC 900  
Db 841 CCAAGACGGCAGCCTTTCGCTACGACATGGAGGTCCAGACTCGGGCATCGACCTGC 900  
QY 901 AGTGCACACTGGCCCTGTATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 901 AGTGCACACTGGCCCTGTATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC 960  
QY 961 TGGAGAACAGCCCTAACCCCTGCCCTCCACCGCCGGCTCCACACTGCGGAAAGCAGCCT 1020  
Db 961 TGGAGAACAGCCCTAACCCCTGCCCTCCACCGCCGGCTCCACACTGCGGAAAGCAGCCT 1020  
QY 1021 TCCTGTCTGGTGCAGATGTGCCCTGAAACACAGAGCTCAGCGTTCACAGGGGATYTG 1080  
Db 1021 TCCTGTCTGGTGCAGATGTGCCCTGAAACACAGAGCTCAGCGTTCACAGGGGATYTG 1080  
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGSCCTCGNCAGACGACGAGTNGGGGG 1140  
Db 1081 NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGSCCTCGNCAGACGACGAGTNGGGGG 1140  
QY 1141 CAAGCNAGAATGCTCCCGAGATTTACANCTGAGCCNTGCCCNACCCCTGCTGAADA 1200  
Db 1141 CAAGCNAGAATGCTCCCGAGATTTACANCTGAGCCNTGCCCNACCCCTGCTGAADA 1200  
QY 1201 AAACAYTNCGCCACGTGAAGAGACAGAGAGATGNCAGGAGTTNNAACCTYGGGGAA 1260  
Db 1201 AAACAYTNCGCCACGTGAAGAGACAGAGAGATGNCAGGAGTTNNAACCTYGGGGAA 1260  
QY 1261 AAAAAACAGGATCTTTNTTCTGCCCTGTCCAGTNCAGTGGCCCTGNACCCGCTTG 1320  
Db 1261 AAAAAACAGGATCTTTNTTCTGCCCTGTCCAGTNCAGTGGCCCTGNACCCGCTTG 1320  
QY 1321 ANTCAGTGACCATTTGTGGCAGANCAGGGGAGAGCAGCTTCAGCGCTGGGTCAAGAGGG 1380

Db 1321 ANTCAGTGACCATTTGTGGCAGANCAGGGGAGAGCAGCTTCAGCGCTGGGTCAAGAGGG 1380  
QY 1381 GTGGCGAGAGCCCTTCGGCCCTCACTCCCTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440  
Db 1381 GTGGCGAGAGCCCTTCGGCCCTCACTCCCTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGTA 1440  
QY 1441 AGGNCNCAANCTCAGGNTTCAAGTGAGAAACAGAGTNCAGCAGGTATGCCCGCCGNTA 1500  
Db 1441 AGGNCNCAANCTCAGGNTTCAAGTGAGAAACAGAGTNCAGCAGGTATGCCCGCCGNTA 1500  
QY 1501 GGTAAANNNGGGGCCCCCTTAAACCCCTTGCTNGGCTNCACCTNGGCCAGCTCANCCC 1560  
Db 1501 GGTAAANNNGGGGCCCCCTTAAACCCCTTGCTNGGCTNCACCTNGGCCAGCTCANCCC 1560  
QY 1561 CTTTTGGGTGTAGGGGAAAAGATGCTGAGCCCTGGGAAAGGCTWCCCTGGTAGAATACAC 1620  
Db 1561 CTTTTGGGTGTAGGGGAAAAGATGCTGAGCCCTGGGAAAGGCTWCCCTGGTAGAATACAC 1620  
QY 1621 CACATTTTTCAGGTTGTGCAACACAGGTCTGTAGTTGACCTCTGGTTCAGCCCAAGGACC 1680  
Db 1621 CACATTTTTCAGGTTGTGCAACACAGGTCTGTAGTTGACCTCTGGTTCAGCCCAAGGACC 1680  
QY 1681 AAGAAGGTGTAAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCTTTTGTGTGTG 1740  
Db 1681 AAGAAGGTGTAAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCTTTTGTGTGTG 1740  
QY 1741 GCTACCACTCTTCCCCAGAGCAGGCCCCGAGCCCCCTTCAGGCCAGCAGCTGCCCCAG 1800  
Db 1741 GCTACCACTCTTCCCCAGAGCAGGCCCCGAGCCCCCTTCAGGCCAGCAGCTGCCCCAG 1800  
QY 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGAGGATATGCTCTGA 1860  
Db 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAGAGTGAAGGTGATGAGGATATGCTCTGA 1860  
QY 1861 CAGAAACAGTCTGTGGATGACATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGAGCGY 1920  
Db 1861 CAGAAACAGTCTGTGGATGACATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGAGCGY 1920  
QY 1921 TCCGGGCCCCCAGNCCCCTACTNATCAGTGTNCCAGCGTCTNGGTTNCCCAGNAGACAC 1980  
Db 1921 TCCGGGCCCCCAGNCCCCTACTNATCAGTGTNCCAGCGTCTNGGTTNCCCAGNAGACAC 1980  
QY 1981 GCTNCAGNCATCANCATGACATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGAGCGY 2040  
Db 1981 GCTNCAGNCATCANCATGACATGATCAGTGTCTTAAAGNAAAGCAGCAGAGAGAGCGY 2040  
QY 2041 CCNACGGCACTTTTGCACTNTGATGNACTCAAGCACTTTTTCATGCTNGCCCTCTNNG 2100  
Db 2041 CCNACGGCACTTTTGCACTNTGATGNACTCAAGCACTTTTTCATGCTNGCCCTCTNNG 2100  
QY 2101 GCAGGNCAGGNCAGGNCAGTGACATCTGTAGNAGCATANGCAAGCCAGAGATGG 2160  
Db 2101 GCAGGNCAGGNCAGGNCAGTGACATCTGTAGNAGCATANGCAAGCCAGAGATGG 2160  
QY 2161 GGTGNAAGGNCACAGTCTTGAGCTGTCCANCATGATGACTNCTCAAACTCTTN 2220  
Db 2161 GGTGNAAGGNCACAGTCTTGAGCTGTCCANCATGATGACTNCTCAAACTCTTN 2220  
QY 2221 NCCAGNATTTCTTAAAGATPAGCANCCTTNCCTCCCACTTGGCCAGCTTAGCTCTTCT 2280  
Db 2221 NCCAGNATTTCTTAAAGATPAGCANCCTTNCCTCCCACTTGGCCAGCTTAGCTCTTCT 2280  
QY 2281 CCAGGGGAGCTANTCAGGACTCAGTAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340  
Db 2281 CCAGGGGAGCTANTCAGGACTCAGTAGCATTAATCAGCTGTGNAATCGTCAGGGGG 2340  
QY 2341 TGTCTGTAGCTCAACCTCTCGGGCAGGGGAGCGGAGACTCCGTTGGGAGAGCTCAT 2400  
Db 2341 TGTCTGTAGCTCAACCTCTCGGGCAGGGGAGCGGAGACTCCGTTGGGAGAGCTCAT 2400  
QY 2401 TCCCATCATCTTGCAGAGAGCAGCTTTNGTGTCCAGCTGTCCACATTTGAGTTCAGCTGCC 2460

Db 2401 TCCACATCTTGGCCAAAGACAGCCTTTNGTCCAGTGTCCACATTCAGTGCAGAGTCTCCC 2460  
Qy 2461 GGGGAGAGAGCCCGGCCCCCAGACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520  
Db 2461 GGGGAGAGAGCCCGGCCCCCAGACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520  
Qy 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGATGACAAAATGTTAAAAAAG 2580  
Db 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGATGACAAAATGTTAAAAAAG 2580  
Qy 2581 GCGTTCGTGGCTCGAATCAAGCTTATCGATACCGTTCGACCTCGAGGGGGG 2631  
Db 2581 GCGTTCGTGGCTCGAATCAAGCTTATCGATACCGTTCGACCTCGAGGGGGG 2631

## RESULT 2

US-09-155-676A-3  
; Sequence 3, Application US/09155676A  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/155,676A  
; FILING DATE: 04-JAN-1999  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IL97/00117  
; FILING DATE: 01-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 117800  
; FILING DATE: 02-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 119133  
; FILING DATE: 26-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25,618  
; REFERENCE/DOCKET NUMBER: WALLACH=21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2631 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-155-676A-3

Query Match 97.1%; Score 2553.8; DB 17; Length 2631;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CCCCTCTCACAGCCCCAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60

1 CCCCTCTCACAGCCCCAGGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60  
61 CTGACGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
61 CTGACGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 180  
121 GCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCCAATCAAGCCAAATTACC 180  
181 ACCAGACCTCCATGCCCCAGCGGAGAGCTTTTCGCAAGGGGCCCAAGGCCCGGGCCAG 240  
181 ACCAGACCTCCATGCCCCAGCGGAGAGCTTTTCGCAAGGGGCCCAAGGCCCGGGCCAG 240  
241 CTGAGGAGACAAACAGGAGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 300  
241 CTGAGGAGACAAACAGGAGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 300  
301 AGCCAAACAAGTCTCTCCCTTGACTTTGACAAAGGAGAGTCTGGGATGTGGGAACCT 360  
301 AGCCAAACAAGTCTCTCCCTTGACTTTGACAAAGGAGAGTCTGGGATGTGGGAACCT 360  
361 TACCTCTGTCTCTCTCTGGAGCGAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAG 420  
361 TACCTCTGTCTCTCTCTGGAGCGAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAG 420  
421 CAACGCTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAATTATTCCTCAACAGCCTGT 480  
421 CAACGCTCCCGAGCAGGAACTGCAGCAGCTGGAAATAGAATTATTCCTCAACAGCCTGT 480  
481 CCCAGCATTCTCTCTGGAGGAGAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 540  
481 CCCAGCATTCTCTCTGGAGGAGCAGGAGCAAAATCTCTCGTGCCTCAGCATCGACAGCC 540  
541 TCTCCCTGTGCGATGACAGTGAAGAAACCATCAAGAGCTCTCAAGCTCGCGGAGCA 600  
541 TCTCCCTGTGCGATGACAGTGAAGAAACCATCAAGAGCTCTCAAGCTCGCGGAGCA 600  
601 CCCTGAGCTCAGGGGTACACTCTTGAGCAGCAGCGGCGGAGCTCGAAGCTCCAGCTGGA 660  
601 CCCTGAGCTCAGGGGTACACTCTTGAGCAGCAGCGGCGGAGCTCGAAGCTCCAGCTGGA 660  
661 ACATGGTGTCTGGCGGGCGGGCCACCCAGACACCCCAAGCTATTTCATGGTGTGAAG 720  
661 ACATGGTGTCTGGCGGGCGGGCCACCCAGACACCCCAAGCTATTTCATGGTGTGAAG 720  
721 TCCAAATAACAGTCTCTTAATGGTGAACACCTGACATCCGGAGTTCCACCGGGTCAAAG 780  
721 TCCAAATAACAGTCTCTTAATGGTGAACACCTGACATCCGGAGTTCCACCGGGTCAAAG 780  
781 TGGGAGACATCGCCACTGGCATCAGACGCGAGATCCAGCTCGAGCCTTCAGCTTGGTCA 840  
781 TGGGAGACATCGCCACTGGCATCAGACGCGAGATCCAGCTCGAGCCTTCAGCTTGGTCA 840  
841 CCAAAGACGGCAGCCTCTTCTGCTACGACATGGAGGTCCAGACTCGGGCATCGACCTGC 900  
841 CCAAAGACGGCAGCCTCTTCTGCTACGACATGGAGGTCCAGACTCGGGCATCGACCTGC 900  
901 AGTGACACTGGCCCCCTGATGGAGCTTCGCTGGAGCTGGAGGTGAAGCATGGCCAGC 960  
901 AGTGACACTGGCCCCCTGATGGAGCTTCGCTGGAGCTGGAGGTGAAGCATGGCCAGC 960  
961 TGGGAAACAGGCCCTTAACCCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 1020  
961 TGGGAAACAGGCCCTTAACCCCTCCACCGCGGCTCCACACTGCCGGAAGCAGCCT 1020  
1021 TCCTCTCGGTGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 TCCTCTCGGTGACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1081 NCCAGCCCCCGGCTCAGCAGTGGGAAACCGAGGGCTCGNAGCAGNAGCAGTGGGGG 1140  
1081 NCCAGCCCCCGGCTCAGCAGTGGGAAACCGAGGGCTCGNAGCAGNAGCAGTGGGGG 1140

Db 1081 NCCAGCCCCCGGCTCARGNTGGAAACAGGGGCTCGNCAGCNAGCNAGGTTNGGGG 1140  
Qy 1141 CAAGNAGAATGCTCCAGGATTTCAACNCTGAGCCCTGCCCCCAGCCCTGCTGAADA 1200  
Db 1141 CAAGNAGAATGCTCCAGGATTTCAACNCTGAGCCCTGCCCCCAGCCCTGCTGAADA 1200  
Qy 1201 AAACAYTNCCCGCAGCTGAAGAGACAGAGAGAGATGNCAGGAGTTNNACCTTGGGAA 1260  
Db 1201 AAACAYTNCCCGCAGCTGAAGAGACAGAGAGAGATGNCAGGAGTTNNACCTTGGGAA 1260  
Qy 1261 AAAAAAGGAGTCTTTNTTCTGCCCCCTGCTCCAGTNGAGTGGCTGNAACCCGCTTGG 1320  
Db 1261 AAAAAAGGAGTCTTTNTTCTGCCCCCTGCTCCAGTNGAGTGGCTGNAACCCGCTTGG 1320  
Qy 1321 ANTCACTGACCATTTGTTGGCAGANCAAGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGGG 1380  
Db 1321 ANTCACTGACCATTTGTTGGCAGANCAAGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGGG 1380  
Qy 1381 GTGGCGAGCCCTTGGGCCCTCACCCTNCCAGGCTGCTGNAAGTGTCAAGTGTGTA 1440  
Db 1381 GTGGCGAGCCCTTGGGCCCTCACCCTNCCAGGCTGCTGNAAGTGTCAAGTGTGTA 1440  
Qy 1441 AGGNCACCAANCTCAGGNTTCAGTGCAAGAACACAGGTCNACAGGTATGCCCGCCGNTA 1500  
Db 1441 AGGNCACCAANCTCAGGNTTCAGTGCAAGAACACAGGTCNACAGGTATGCCCGCCGNTA 1500  
Qy 1501 GGTAAANNNGGGGGCCCTCTNAACCCCTTGGCTNGGCCCTNACCTNGGCCAGCTCANCCC 1560  
Db 1501 GGTAAANNNGGGGGCCCTCTNAACCCCTTGGCTNGGCCCTNACCTNGGCCAGCTCANCCC 1560  
Qy 1561 CTTTGGGTGAGGGGAAAGATGCTGACCCCTGGGAGGCTWCCCTGGTGAATACAC 1620  
Db 1561 CTTTGGGTGAGGGGAAAGATGCTGACCCCTGGGAGGCTWCCCTGGTGAATACAC 1620  
Qy 1621 CACACTTTTCAAGTGTGTTGCAACACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC 1680  
Db 1621 CACACTTTTCAAGTGTGTTGCAACACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC 1680  
Qy 1681 AAAGAAGTGTGTAAGTGAAGTGTCTCAGTNCACCCAGACATGTCGCCCTTTGCTGCTG 1740  
Db 1681 AAAGAAGTGTGTAAGTGAAGTGTCTCAGTNCACCCAGACATGTCGCCCTTTGCTGCTG 1740  
Qy 1741 GCTACCACTTCTCCAGAGCAGAGGCCCCGAGCCCTTCAAGGCCAGCAGCTGCCCCAG 1800  
Db 1741 GCTACCACTTCTCCAGAGCAGAGGCCCCGAGCCCTTCAAGGCCAGCAGCTGCCCCAG 1800  
Qy 1801 ACTGCTGGCACTCAGTTCCTCTCATCTGAAGGTGAAGGTGATGAGGATATGCCCTGA 1860  
Db 1801 ACTGCTGGCACTCAGTTCCTCTCATCTGAAGGTGAAGGTGATGAGGATATGCCCTGA 1860  
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGGNAAGCAGCAGAGAGAGCY 1920  
Db 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGGNAAGCAGCAGAGAGAGCY 1920  
Qy 1921 TCCGGCGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTGTTNCCCGCAGNAGCACA 1980  
Db 1921 TCCGGCGCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTGTTNCCCGCAGNAGCACA 1980  
Qy 1981 GCTNAGNCATCANCACTGACACTNCCCTNGCCCTGCCCCCTGGCCGAGGAGGATG 2040  
Db 1981 GCTNAGNCATCANCACTGACACTNCCCTNGCCCTGCCCCCTGGCCGAGGAGGATG 2040  
Qy 2041 CCGNAGGCACTTGGACNTCTGATGACCTCAAGGACCTTTCATGCTNGCCCTCTNNG 2100  
Db 2041 CCGNAGGCACTTGGACNTCTGATGACCTTCAAGGACCTTTCATGCTNGCCCTCTNNG 2100  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATFANGCAAGCCAGGAGATGG 2160  
Db 2101 GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATFANGCAAGCCAGGAGATGG 2160  
Qy 2161 GGTGNAAGGNCACAGCTTTGAGTGTCCANCATGATGATGATGATGATGATGATGATGATG 2220  
Db 2161 GGTGNAAGGNCACAGCTTTGAGTGTCCANCATGATGATGATGATGATGATGATGATGATG 2220

## RESULT 3

US-09-155-676-6

; Sequence 6, Application US/09155676

; GENERAL INFORMATION:

; APPLICANT: WALLACH, David

; APPLICANT: MALININ, Nikolai

; APPLICANT: BOLDIN, Mark

; APPLICANT: KOVALENKO, Andrei

; APPLICANT: METT, Igor

; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED

; FACTOR (TRAF), THEIR PREPARATION AND USE

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.

; STREET: 624 Ninth Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/155,676

; FILING DATE: 04-JAN-1999

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/IL97/00117

; FILING DATE: 01-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: IL 117800

; FILING DATE: 02-APR-1996

; APPLICATION NUMBER: IL 119133

; FILING DATE: 26-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: BROWDY, Roger L.

; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: WALLACH-21

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-628-5197

Qy 2221 NCCAGNATTTCTTAAGAAATAGCANCCCCCTTNNCCCAATGCCCCAGCTTAGCCTCTTCT 2280  
Db 2221 NCCAGNATTTCTTAAGAAATAGCANCCCCCTTNNCCCAATGCCCCAGCTTAGCCTCTTCT 2280  
Qy 2281 CCCAGGGAGCTANTCAGGACTCAGCTAGCATTAATCAGCTGTGNAATGTCAGGGG 2340  
Db 2281 CCCAGGGAGCTANTCAGGACTCAGCTAGCATTAATCAGCTGTGNAATGTCAGGGG 2340  
Qy 2341 TGTCTGCTAGCTCAACCTCTCTGGGCGAGGGAGCCGCGAGACTCCGTGGGAGAGAGCTCAT 2400  
Db 2341 TGTCTGCTAGCTCAACCTCTCTGGGCGAGGGAGCCGCGAGACTCCGTGGGAGAGAGCTCAT 2400  
Qy 2401 TCCACATCTTGGCAAGACAGCCTTTTNGTCCAGCTGTCCATTAATGATGAGCTGCTGCC 2460  
Db 2401 TCCACATCTTGGCAAGACAGCCTTTTNGTCCAGCTGTCCATTAATGATGAGCTGCTGCC 2460  
Qy 2461 GGGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520  
Db 2461 GGGGAGAGAGCCCCGGCCCCCAGCACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTG 2520  
Qy 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGCTGATGACAAATGTTAAAAAAG 2580  
Db 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGCTGATGACAAATGTTAAAAAAG 2580  
Qy 2581 GCCTTCGTGGCTCGAATCAAGCTTTATCGATACCGTCGACCTCGAGGGGG 2631  
Db 2581 GCCTTCGTGGCTCGAATCAAGCTTTATCGATACCGTCGACCTCGAGGGGG 2631

TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4596 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-155-676-6

Query Match 70.4%; Score 1851.2; DB 17; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCAGCCATCCAAAGAGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60  
Db CCCCTCTCACAGCCAGCCATCCAAAGAGGGCTGAGAAAGAGCCCATCCACCGCGTGT 2158

QY 61 CTGACGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGGAGTCTGAAGA 120  
Db CTGACGGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGGAGTCTGAAGA 2218

QY 121 GCCCTTGGAGGGAGATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 180  
Db GCCCTTGGAGGGAGATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC 2278

QY 181 ACCAGACCTTCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCCGGCCAG 240  
Db ACCAGACCTTCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCCGGCCAG 2338

QY 241 CTGAGAGACAAGCAGCAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 300  
Db CTGAGAGACAAGCAGCAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCCAG 2398

QY 301 AGCCAAACAACTCTCCCTCTGACTTTGACAAAGAGAGAGTCTGGGATGTGGGAACCT 360  
Db AGCCAAACAACTCTCCCTCTGACTTTGACAAAGAGAGAGTCTGGGATGTGGGAACCT 2458

QY 361 TACCTCTGCTCCCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCACCAGAGCGGAAG 420  
Db TACCTCTGCTCCCTGGAGCCAGCCCTTGCCAGAAACCCAGCTCACCAGAGCGGAAG 2518

QY 421 CAACCTCCCGAGCAGGAACTCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 480  
Db CAACCTCCCGAGCAGGAACTCAGCAGCTGGAATAGAAATATTCTCAACAGCCTGT 2578

QY 481 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATCTCTGCTGCTTCAGCATCGACGCC 540  
Db CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATCTCTGCTGCTTCAGCATCGACGCC 2638

QY 541 TCTCCCTGTCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACA 600  
Db TCTCCCTGTCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACA 2698

QY 601 CCCTGAGCTCAGGCTACATCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTCCAGTTGA 660  
Db CCCTGAGCTCAGGCTACATCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTCCAGTTGA 2758

QY 661 ACATGGTGTGGCCGGGGCGGCCACCAAGACCCCAAGCTATTTCAATGGTGTGAAG 720  
Db ACATGGTGTGGCCGGGGCGGCCACCAAGACCCCAAGCTATTTCAATGGTGTGAAG 2818

QY 721 TCCAAATACAGTCTCTTAATGGTGAACCTGACATCCGGGAGTTCCACCGGGTCAAAG 780  
Db TCCAAATACAGTCTCTTAATGGTGAACCTGACATCCGGGAGTTCCACCGGGTCAAAG 2878

QY 781 TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTCAGCTTGGTCA 840  
Db TGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTCAGCTTGGTCA 2938

QY 841 CCAAGAGCGGCAGCTGTTCCGTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGC 900  
Db CCAAGAGCGGCAGCTGTTCCGTACGACATGGAGGTGCCAGCTCGGGCATCGACCTGC 2998

QY 901 AGTGACACTGGCCCTGATGTCAGCTTCGCTGGAGCTGGAGGGTCAAGCATGCGCAC 960  
Db AGTGACACTGGCCCTGATGTCAGCTTCGCTGGAGCTGGAGGGTCAAGCATGCGCAC 3058

QY 961 TGGAGAACAGGGCCCTAACCTCCACCGCGGCTCCACACTCCCGGAAGCAGCT 1020  
Db TGGAGAACAGGGCCCTAACCTCCACCGCGGCTCCACACTCCCGGAAGCAGCT 3118

QY 1021 TCTGCTCGGTGACCATGCTGCTGCTGAAACACACAGGCTCAGCGCTTCCAGGGGATY 1080  
Db TCTGCTCGGTGACCATGCTGCTGCTGAAACACACAGGCTCAGCGCTTCCAGGGGATY 3176

QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCTCGNCAGCNAGNAAGTNGGGG 1140  
Db NCCAGCCCCCGGCTCARCAGNTGGGAACACAGGCTCGNCAGCNAGNAAGTNGGGG 3229

QY 1141 CAAGCAGAAATGCTCCAGGATTTACANCTGAGCCCTGCCANCCCTGCTCAADA 1200  
Db CAAGCAGAAATGCTCCAGGATTTACANCTGAGCCCTGCCANCCCTGCTCAADA 3283

QY 1201 AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGGAGTTNNACCTYGGGAA 1260  
Db AAACAYTNCGCCACGCTGAAGAGACAGAGGAGATGNCAGGAGTTNNACCTYGGGAA 3339

QY 1261 AAAAAACAGGGATCTTTTCTGCTCCCTGCTCCAGTNCAGTGGGCTGNAACCGCTTG 1320  
Db AAAAAACAGGGATCTTTTCTGCTCCCTGCTCCAGTNCAGTGGGCTGNAACCGCTTG 3395

QY 1321 ANTCACTGACATTTGTTGGCAGANAGGGAGAGCAGCTCCAGCCTGGGTGAGAAGG 1380  
Db ANTCACTGACATTTGTTGGCAGANAGGGAGAGCAGCTCCAGCCTGGGTGAGAAGG 3454

QY 1381 GTGGGCGAGCCTTCGGCCCTCACCCTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA 1440  
Db GTGGGCGAGCCTTCGGCCCTCACCCTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA 3512

QY 1441 AGGNCCCAAANCTCAGNTTCAGTGCAAGAACAGTNCAGCAGGATGCGCCGCCGNTA 1500  
Db AGGNCCCAAANCTCAGNTTCAGTGCAAGAACAGTNCAGCAGGATGCGCCGCCGNTA 3568

QY 1501 GGTAAANNNGGGGCTCTTNAACCCCTTGCTNGGCTNCACCTNGGCGCAGCTCANCCC 1560  
Db GGTAAANNNGGGGCTCTTNAACCCCTTGCTNGGCTNCACCTNGGCGCAGCTCANCCC 3620

QY 1561 CTTTGGGTGAGGGGAAAAGATGCTGACCCCTGGGAAGGCTWCCCTGTFAGAAATAC 1620  
Db CTTTGGGTGAGGGGAAAAGATGCTGACCCCTGGGAAGGCTWCCCTGTFAGAAATAC 3679

QY 1621 CACACTTTTCAGGTTGTAAGTGAAGTGGTTCTCAGTNCAGTCCAGCATATGTGCTGT 1680  
Db CACACTTTTCAGGTTGTAAGTGAAGTGGTTCTCAGTNCAGTCCAGCATATGTGCTGT 3739

QY 1681 AAAGAAAGTGTGTAAGTGAAGTGGTTCTCAGTNCAGTCCAGCATATGTGCTGT 1740  
Db AAAGAAAGTGTGTAAGTGAAGTGGTTCTCAGTNCAGTCCAGCATATGTGCTGT 3798

QY 1741 GCTACACTCTTCCCAGAGCAGAGGCCCGAGCCCTTCAGGCCAGCACTGCCCCAG 1800  
Db GCTACACTCTTCCCAGAGCAGAGGCCCGAGCCCTTCAGGCCAGCACTGCCCCAG 3858

QY 1801 ACTCGCTGCACTCAGTTCCCTCATCTCTAAAGGTGAAGGTGATGACAGGATATGCTGA 1860  
Db ACTCGCTGCACTCAGTTCCCTCATCTCTAAAGGTGAAGGTGATGACAGGATATGCTGA 3918

QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGAG 1920  
Db CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGAG 3975

QY 1921 TCCGGCGCCCAAGNCCCACTNATCAGTGTNCCAGGCTGCTNGTTNCCCAAGNAGCA 1980  
Db TCCGGCGCCCAAGNCCCACTNATCAGTGTNCCAGGCTGCTNGTTNCCCAAGNAGCA 4031



QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGCTCAAAG 780  
Db 2819 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGCTCAAAG 2878  
QY 781 TGGGAGACATCGCCACTGGGATCGAGCAGCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 840  
Db 2879 TGGGAGACATCGCCACTGGGATCGAGCAGCAGATCCAGCTGCGAGCTTCAGCTTGGTCA 2938  
QY 841 CCAAGACGGGACGCTCTCGCTACGACATGAGGTGCCAGATCGGGCATCGACCTGC 900  
Db 2939 CCAAGACGGGACGCTCTCGCTACGACATGAGGTGCCAGATCGGGCATCGACCTGC 2998  
QY 901 AGTGACACTGGCCCTCATGCGAGCTTCGGCTGGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 2999 AGTGACACTGGCCCTCATGCGAGCTTCGGCTGGAGCTGGAGGTCAAGCATGGCCAGC 3058  
QY 961 TGGAGAACAGGCTTAAACCTGCTCCACCGCGGGTCCACACTGCGGGAAGAGAGCT 1020  
Db 3059 TGGAGAACAGGCTTAAACCTGCTCCACCGCGGGTCCACACTGCGGGAAGAGAGCT 3118  
QY 1021 TCCTGCTCGGTGACAGATGCTGCTGAAACACAGAGCTCAGCGTTCGCCAGGGGATVTG 1080  
Db 3119 TCCTGCTCGGTGACAGATGCTGCTGAAACACAGAGCTCAGCGTTCGCCAGGGGATVTG 3176  
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACCCAGGGCTTCGNCAGCAGNAGNAAAGTNGGGG 1140  
Db 3177 GCCAGCCCCCGGCTCA--CAGTGGGAACAGGGCTTCG-----CAGCAGCAAGGTGGG 3229  
QY 1141 CAAGNAGATGCTCCAGAGATTTACANCTGAGCCCTGAGCCCTGCCAGCCCTGCTGAADA 1200  
Db 3230 GCAAGCAAGATGCTCCAGAGATTTACANCTGAGCCCTGAGCCCTGCCAGCCCTGCTGA 3283  
QY 1201 AAACATYNNCCGACGTGAAGAGACAGAGAGAGATGNCAGGAGTTNNACTCTGGGGA 1260  
Db 3284 AAAAATCCCGCCAGTGAAGAGACAGAGAGAGATGNCAGGAGTTNNACTCTGGGGA 3339  
QY 1261 ACAAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGTGTGCGCTGNCACCCGCTTGG 1320  
Db 3340 ACAAAACAGGATCTTTTTCGCCCCCTGCTCCAGTNCAGTGTGCGCTGNCACCCGCTTGG 3395  
QY 1321 ANTACGTACATTTGTTGGGAGANCAGGGGAGAGAGCTTCCAGCCCTGGGTGAGAAAGG 1380  
Db 3396 GATCAGTGACCATTTGTTGGCAGA-CAGGGGAGAGAGCTTCCAGCCCTGGGTGAGAAAGG 3454  
QY 1381 GTGGCGAGCCCTTCGGCCCTTCAACCTNCCAGGCTGCTGTGNCAGTGTCAAGTGTGTA 1440  
Db 3455 GTGGCGAGCCCTTCGGCCCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAAC 3512  
QY 1441 AGGGNCCCAANCTCAGGNTTCAGTGTGAGAACACAGGTNACAGAGTATGCCCGCCGNTA 1500  
Db 3513 AGGGNCCCAANCTCAGGNTTCAGTGTGAGAACACAGGTNACAGAGTATGCCCGCCGNTA 3568  
QY 1501 GGTAAANGGGGCTCTNAAACCCCTTGCCTNCGCCCTNACCTNCGCCAGCTCANCC 1560  
Db 3569 GTTAA-----GGGGGCTCTNAAACCCCTTGCCTNCGCCCTNACCTNCGCCAGCTCA-CC 3620  
QY 1561 CTTTTGGGTGAGGGAAGAATGCTGACCTCGGGAAGGTGCTGCTGAGTAAATACAC 1620  
Db 3621 CTTTTGGGTGAGGGAAGAATGCTGACCTCGGGAAGGTGCTGCTGAGTAAATACAC 3679  
QY 1621 CACACTTTTTCAGGTTGTTGCAACACAGGTCTGAGTGTGACCTCTGTTGAGCAGCAAGGACC 1680  
Db 3680 CACACTTTTTCAGGTTGTTGCAACACAGGTCTGAGTGTGACCTCTGTTGAGCAGCAAGGACC 3739  
QY 1681 AAAGAGGTGCTAGTGAAGTGTCTCAGTNCAGATGTCAGATGTCGCTTGTGCTG 1740  
Db 3740 AAAGAGGTGCTAGTGAAGTGTCTCAGTNCAGATGTCAGATGTCGCTTGTGCTG 3798  
QY 1741 GCTACCACTCTTCCCGAGAGCAGCGCCCGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 1800  
Db 3799 GCTACCACTCTTCCCGAGAGCAGCGCCCGAGCCCTTTCAGGCCCCAGCACTGCCCCAG 3858  
QY 1801 ACTCGTGGCACTCAGTTCCTCATCTGTAAGAGGTGAAGGGTGTGAGGATATGCTCTGA 1860

Db 3859 ACTCGTGGCACTCAGTTCCTCATCTGTAAGGTGAAGGTGTATGAGGATATGCTCTGA 3918  
QY 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGAGC 1920  
Db 3919 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGAGAGC 3975  
QY 1921 TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTNCCAGNAGCAGC 1980  
Db 3976 TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTNCCAGNAGCAGC 4031  
QY 1981 GCTNCAAGCATCANCACTGACATNCACTTNGCCCTGCCCCCTNCCAGNAGAGAGTCTG 2040  
Db 4032 TCCGGCCCCCAGNCCCACTNATCAGTGTNCCAGCGTGTCTNCCAGNAGAGAGTCTG 4082  
QY 2041 CCGNACCGCACTTTCGACATCTGATGATGATGATGATGATGATGATGATGATGATGATG 2100  
Db 4083 CCG-ACGGCACTTTCGACATCTGATGATGATGATGATGATGATGATGATGATGATGATG 4139  
QY 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAANGCCAGAGATGG 2160  
Db 4140 GGGCAGGCGAGGCGAGTGACACT-----GTAGGAGCATAGCAAGCCAGAGATGG 4189  
QY 2161 GGTNAAAGGAGNACAGTCTTGTAGCTGTCCANCATGATGATGATGATGATGATGATGATG 2220  
Db 4190 GGTGAA--GGGACACAGTCTTGTAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4244  
QY 2221 NCCAGNATTTCTTAAAGATAGCANCCCTTNCCTTCCCATTTGCCCGAGCTTAGCCTTCT 2280  
Db 4245 --CCAGATTTCTTAAAGATAGCACCCCTT--TTCCCATTTGCCCGAGCTTAGCCTTCT 4300  
QY 2281 CCCAGGGAGCTNCTCAGGACTCAGCTAGCATTAATTAATTAATTAATTAATTAATTAAT 2340  
Db 4301 CCCAGGGAGCTA-CTCAGGACTCAGCTAGCATTAATTAATTAATTAATTAATTAATTAAT 4358  
QY 2341 TGTCTGTAGCTCAACCTCTCTGGGCGAGGAGCGCGAGACTCGGTGGGAGAGCTCAT 2400  
Db 4359 TGTCTGTAGCTCAACCTCTCTGGGCGAGGAGCGCGAGACTCGGTGGGAGAGCTCAT 4418  
QY 2401 TCCACATCTTGCACAGAGAGCTTNGTCCAGCTGTCCACATTTAGTTCAGACTGCTCCC 2460  
Db 4419 TCCACATCTTGCACAGAGAGCTTNGTCCAGCTGTCCACATTTAGTTCAGACTGCTCCC 4477  
QY 2461 GGGGAGAGAGCCCCCGCCCCCAGACATATAAGAACTGAGCCTTGGTACTGAGAGTCTG 2520  
Db 4478 GGGGAGAGAGCCCCCGCCCCCAGACATATAAGAACTGAGCCTTGGTACTGAGAGTCTG 4537  
QY 2521 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGATGATGATGATGATGATG 2579  
Db 4538 GGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTGGGGTGTGATGATGATGATGATGATG 4596

RESULT 5  
PCT-US01-08631-15294  
; Sequence 15294, Application PC/TUS0108631  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-049  
; CURRENT APPLICATION NUMBER: PCT/US01/08631  
; PCT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540, 217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649, 167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 15294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR

LOCATION: (1117)..(1581)  
OTHER INFORMATION: 100% homologous to Homo sapiens sperm protein 10, accession  
; OTHER INFORMATION: number M82968, Smith-Waterman Score=809.  
PCT-US01-08631-15294

	Query Match	70.3%	Score 1850.2;	DB 1;	Length 4596;
	Best Local Similarity	90.5%	Pred. No. 0;		
	Matches 2334;	Conservative 4;	Mismatches 159;	Indels 81;	Gaps 34;
Qy	1	CCCTCTCACAGCCAGCCAGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT	60		
Db	2100	CCCTCTCACAGCCAGCCAGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGT	2159		
Qy	61	CTGAGGCGAGCTGGAGGAGAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA	120		
Db	2160	CTGAGGCGAGCTGGAGGAGAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA	2219		
Qy	121	GCCCTTGGAGGGGAGATATAAGCAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	180		
Db	2220	GCCCTTGGAGGGGAGATATAAGCAACCAAGACATCCACCGCCAAATCAAGCCAAATTACC	2279		
Qy	181	ACAGAGCCCTCCATGCCAGCCAGAGAGGCTTTCCGCAAGGGGCCCCAGGGCCCGGCCAG	240		
Db	2280	ACAGAGCCCTCCATGCCAGCCAGAGAGGCTTTCCGCAAGGGGCCCCAGGGCCCGGCCAG	2339		
Qy	241	CTGAGGAGACAACAGGAGAGCCCTTAAGTCTCAGCTCTCTCTCCACAGAGCCCGCCAG	300		
Db	2340	CTGAGGAGACAACAGGAGAGCCCTTAAGTCTCAGCTCTCTCTCCACAGAGCCCGCCAG	2399		
Qy	301	AGCCAAACAGTCTCCTCCCTTGAATTTGACGAGAGGAGTCTGGATGTGGAAACCT	360		
Db	2400	AGCCAAACAGTCTCCTCCCTTGAATTTGACGAGAGGAGTCTGGATGTGGAAACCT	2459		
Qy	361	TACCTCTGTCTCCTCGAGCAGCCCTCTGCAGAAACCCAGCTCACACAGAGCGGAAAG	420		
Db	2460	TACCTCTGTCTCCTCGAGCAGCCCTCTGCAGAAACCCAGCTCACACAGAGCGGAAAG	2519		
Qy	421	CAACCGTCCGGAGAGAGAACTGCAGAGCTGGAAATAGAAATATTCTCAACAGCCCTGT	480		
Db	2520	CAACCGTCCGGAGAGAGAACTGCAGAGCTGGAAATAGAAATATTCTCAACAGCCCTGT	2579		
Qy	481	CCAGCCATTTCTCTGGAGGAGCAGGAGCAATTTCTCTGTCCTCAGCATGCAGAGCC	540		
Db	2580	CCAGCCATTTCTCTGGAGGAGCAGGAGCAATTTCTCTGTCCTCAGCATGCAGAGCC	2639		
Qy	541	TCCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGCTCCGGGACA	600		
Db	2640	TCCTCCCTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGCTCCGGGACA	2699		
Qy	601	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA	660		
Db	2700	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA	2759		
Qy	661	ACATGTGTGCGCGGGGGGGCCCAACGACACCCCAAGCTATTCAATGGTGTGAAG	720		
Db	2760	ACATGTGTGCGCGGGGGGGGGCCCAACGACACCCCAAGCTATTCAATGGTGTGAAG	2819		
Qy	721	TCCAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGGTCAAG	780		
Db	2820	TCCAATACAGTCTCTTAATGGTGAACACCTGACATCCCGGAGTTCACCGGGTCAAG	2879		
Qy	781	TGGGAGACATCGCCATCGGATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTCA	840		
Db	2880	TGGGAGACATCGCCATCGGATCAGCAGCCAGATCCAGCTGCAGCTTCAGCTTGGTCA	2939		
Qy	841	CCAAAGACGGGAGCTGTTCGTTGCAACATGAGAGTGCAGACTCGGGCATCGACTGC	900		
Db	2940	CCAAAGACGGGAGCTGTTCGTTGCAACATGAGAGTGCAGACTCGGGCATCGACTGC	2999		
Qy	901	AGTGCACTGGCCCTGATGGAGCTTCGCTGAGCTGGAGGTCAGCATGGCCAGC	960		
Db	3000	AGTGCACTGGCCCTGATGGAGCTTCGCTGAGCTGGAGGTCAGCATGGCCAGC	3059		

Qy	961	TGAGAAACAGGCGCTTAACCTCGCCCTCCACGCGGGCTCCACACTGCGGAAAGACGCT	1020
Db	3060	TGAGAAACAGGCGCTTAACCTCGCCCTCCACGCGGGCTCCACACTGCGGAAAGACGCT	3119
Qy	1021	TCCTGTCTCGGTGACGATGCTGCCCTGAAAAACACAGCTCAGCGTTCCAGGGGATYTG	1080
Db	3120	TCCTGTCTCGGTGACGATGCTGCCCTGAAAAACACAGCTCAGCGTTCCAGGGGATY--	3177
Qy	1081	NCAGAGCCCGCGCTCARCAGNTGGGAACACAGGCGCTCGNACAGNAGCTTNGGGG	1140
Db	3178	GCAGAGCCCGCGCTCA--CAGTGGGAACACAGGCGCTCG-----CAGCAGAGGTGGG	3230
Qy	1141	CAAGCAGAGTGTCTCCAGGATTTACANCTGAGCCNTGCCANCCCTGCTGAADA	1200
Db	3231	GCAAGCAGAGTGTCTCCAGGATTTACACCTGAGCCCTGCCAC-----CCTGCTGA	3284
Qy	1201	AAACATNCCGCGCAGTGAAGACAGAGAGGATGAGGATGAGGATGAGGATGAGGAA	1260
Db	3285	AAACATNCCGCGCAGTGAAGACAGAGAGGATGAGGATGAGGATGAGGATGAGGAA	3340
Qy	1261	ACAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGATTTGGCCCTGNACCCGCTGG	1320
Db	3341	ACAAACAGGATCTTTTCTGCCCCCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG	3396
Qy	1321	ANTCAGTGACATTTGTGGCAGNACAGGGAGAGAGCTTCCAGCTGGGTGAGAGGG	1380
Db	3397	GATCAGTGACATTTGTGGCAG--CAGGGAGAGAGCTTCCAGCTGGGTGAGAGGG	3455
Qy	1381	GTGGGAGGCGCTTCCGCCCCCTCACCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA	1440
Db	3456	GTGGGAGGCGCTTCCGCCCCCTCACCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA	3513
Qy	1441	AGGNCCCAAANTCTCAGGNTTCACTGCAGAACACAGGTCNACAGAGTATGCGGCCGNTA	1500
Db	3514	AGGNCCCAAATCTC--AGGTTCACTGCAGAACACAGT--CAGCAGTATGCGGCCGNTAG	3569
Qy	1501	GGTTAANGGGGCGCTCTNAAACCCCTTGGCTNCGCTNCACTNGGCGAGCTCANCCC	1560
Db	3570	GTAA-----GGGGCGCTCTTAAACCCCTTGGCTGCGCTCA--CCTGGCCAGCTCA--CCC	3621
Qy	1561	CTTTTGGGTGAGGGGAAAGATGCTGACCTGGGAGGCTTCCCTGTGATGATACAC	1620
Db	3622	CTTTTGGGTGAGGGGAAAGATGCTGACCTTGGGAGGCT--CCTGTGATGATACAC	3680
Qy	1621	CACATTTTTCAGGTGTGCAACACAGGCTCTCAGTTGACCTCTGTTTTCAGCAGGAC	1680
Db	3681	CACATTTTTCAGGTGTGCAACACAGGCTCTCAGTTGACCTCTGTTTTCAGCAGGAC	3740
Qy	1681	AAAGAGGTGTGTAAGTGAAGTGTCTCAGTNCCTCCAGACATGTGCCCTTTGCTGTG	1740
Db	3741	AAAGAGGTGTGTAAGTGAAGTGTCTCAGT--CCCAGACATGTGCCCTTTGCTGTG	3799
Qy	1741	GCTACACCTTTCCTCCAGAGCAGCGCCCGAGCCCTTTCAGGCGCAGCAGCTGCCCCAG	1800
Db	3800	GCTACACCTTTCCTCCAGAGCAGCGCCCGAGCCCTTTCAGGCGCAGCAGCTGCCCCAG	3859
Qy	1801	ACTCGGTGCGCTCAGTTCCTCATCTGTAAAGTGAAGGTCATGAGGATATGCTGA	1860
Db	3860	ACTCGGTGCGCTCAGTTCCTCATCTGTAAAGTGAAGGTCATGAGGATATGCTGA	3919
Qy	1861	CAGGAAACAGTGTGTGATGGACATGATGATGATGATGATGATGATGATGATGATGATG	1920
Db	3920	CAGGAAACAGTGTGTGATGGACATGATGATGATGATGATGATGATGATGATGATGATG	3976
Qy	1921	TCCGCGCCCGCAGNCCCACTTATCAGTGTNCCAGGTCGCTGCTGCTGCTGCTGCTGCTG	1980
Db	3977	TCCGCGCCCGCAGNCCCACTTATCAGTGTNCCAGGTCGCTGCTGCTGCTGCTGCTGCTG	4032
Qy	1981	GCTNACAGNATCANCACTGACCTNCCAGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040
Db	4033	-----TAGCATATCATGACCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4083
Qy	2041	CCGACGCGCTTTTGACNTCTGATGACCTCAAGACCTTTTCTGCTGCTGCTGCTGCTGCTG	2100



```

Db 4084 CCG-ACGGCACTTGGCACT--CTGATGACCTAAAGACATTTTCATGGTGCCTCTCGCA 4140
Qy 2101 CGAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATANGAAGCCAGAGATGG 2160
Db 4141 GGGAGGCGGAGGAGTGCACCT-----GTAGGAGCATAGCAAGCCAGAGATGG 4190
Qy 2161 GGTGNAAGGGANCAAGCTTTGAGCTGCCANCATGCAATGACATNCTCAAACTCTTN 2220
Db 4191 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT- 4245
Qy 2221 NCCAGNATTTCTTAAGAAATAGCANCCTTTNCCCAATTCGCCAGCTTAGCTCTTCT 2280
Db 4246 --CCAGATTTCTTAAGAAATAGCACCCCT--TTCCCAATTCGCCAGCTTAGCTCTTCT 4301
Qy 2281 CCCAGGAGCTANTCAGGACTCAGCTAGCATTAATACATGCTGNAATCGTCAGGGGG 2340
Db 4302 CCCAGGAGCTA-CTCAGGACTCAGCTAGCATTAATACATGCTG- AATCGTCAGGGGG 4359
Qy 2341 TGTCTGTAGCTCAACCTCTCTGGGCGAGGAGCGCCGAGACTCCGTGGGAGAAAGCTCAT 2400
Db 4360 TGTCTGTAGCTCAACCTCTCTGGGCGAGGAGCGCCGAGACTCCGTGGGAGAAAGCTCAT 4419
Qy 2401 TCCCAATCTTGGCCAAAGACAGCCTTTNGTCCAGCTGTCCAATTGAGTCAGCTGCC 2460
Db 4420 TCCCAATCTTGGCCAAAGACAGCCTTT- GTCCAGCTGTCCAATTGAGTCAGCTGCC 4478
Qy 2461 GGGGAGAGAGCCCGGCGCCGAGCATAAAGAACTGAGCTTGGTACTCGAGAGCTG 2520
Db 4479 GGGGAGAGAGCCCGGCGCCGAGCATAAAGAACTGAGCTTGGTACTCGAGAGCTG 4538
Qy 2521 GGTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 2578
Db 4539 GGTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 4596

RESULT 6
PCT-US02-18947-975
; Sequence 975, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; TYPE: DNA
; LENGTH: 4596
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-975

Query Match 70.3%; Score 1850.2; DB 1; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCTCTCAGCGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60
Db 2100 CCCTCTCAGCGCCAGGCGCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159
Qy 61 CTGACGCGAGCTGGGAGGAGGTGAACCGGCGCACTACAGCAATGGGAGGTCTGAAGA 120
Db 2160 CTGACGCGAGCTGGGAGGAGGTGAACCGGCGCACTACAGCAATGGGAGGTCTGAAGA 2219
Qy 121 GCCCTTGGGGGAGAAATATAAGAACCAAGACATCCACGCCAAATCAAGCAATACC 180
Db 2220 GCCCTTGGGGGAGAAATATAAGAACCAAGACATCCACGCCAAATCAAGCAATACC 2279
```

```

Qy 181 ACCAGACCTTCATGCCCCAGCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCCGGCAG 240
Db 2280 ACCAGACCTTCATGCCCCAGCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCCGGCAG 2339
Qy 241 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTCCAGAGCTCTCTCCACAGAGAGCCCGCAG 300
Db 2340 CTGAGGAGACAAAGGAGAGAGCCCTTAAGCTCCAGAGCTCTCTCCACAGAGAGCCCGCAG 2399
Qy 301 AGCCAAACAAAGTCTCTCTCCCTTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGGAAACCTT 360
Db 2400 AGCCAAACAAAGTCTCTCTCCCTTCTGACTTTGAGCAAGGAGGAGTCTGGATGTGGGAAACCTT 2459
Qy 361 TACTCTGTCTCTCTCTGGAGCGAGCCCTCTCCAGAAACCCAGCTCACCAGAGAGGGAAG 420
Db 2460 TACTCTGTCTCTCTCTGGAGCGAGCCCTCTCCAGAAACCCAGCTCACCAGAGAGGGAAG 2519
Qy 421 CAACCGTCCCGGAGCAGGAACTCGACAGCTGGAATATAGAAATATTCTCTCAACAGCCTGT 480
Db 2520 CAACCGTCCCGGAGCAGGAACTCGACAGCTGGAATATAGAAATATTCTCTCAACAGCCTGT 2579
Qy 481 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 540
Db 2580 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 2639
Qy 541 TCTCCCTGTCTGGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTTCGCGGACA 600
Db 2640 TCTCCCTGTCTGGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTTCGCGGACA 2699
Qy 601 CCCTGAGCTCAGGCGTACACTCTTGAGAGCAGCAGGCGGAGGCTCGAAGCTTCAAGCTGGA 660
Db 2700 CCCTGAGCTCAGGCGTACACTCTTGAGAGCAGCAGGCGGAGGCTCGAAGCTTCAAGCTGGA 2759
Qy 661 ACATGTGTCTGGCCCGGGGGCGGCCCAAGCAACCCCAAGCTATTTCAATGGTGTGAAAG 720
Db 2760 ACATGTGTCTGGCCCGGGGGCGGCCCAAGCAACCCCAAGCTATTTCAATGGTGTGAAAG 2819
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACTCCGGGAGTTCCACGGGTCAAAG 780
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGCACTCCGGGAGTTCCACGGGTCAAAG 2879
Qy 781 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCGAGCTTCAAGCTTGTGCA 840
Db 2880 TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCGAGCTTCAAGCTTGTGCA 2939
Qy 841 CCAAAGAGCGGCGAGCTTCTGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 900
Db 2940 CCAAAGAGCGGCGAGCTTCTGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGC 2999
Qy 901 AGTGACACTGGCCCTGTATGGGAGCTTCCGCTGGAGCTGGAGGCTCAAGCATGGCCAGC 960
Db 3000 AGTGACACTGGCCCTGTATGGGAGCTTCCGCTGGAGCTGGAGGCTCAAGCATGGCCAGC 3059
Qy 961 TGGAGAACAGGCGCTTAACCTGCGCTCCACCGCGGCTTCCACACTGCGCGGAAAGCAGCT 1020
Db 3060 TGGAGAACAGGCGCTTAACCTGCGCTCCACCGCGGCTTCCACACTGCGCGGAAAGCAGCT 3119
Qy 1021 TCCTGTCTGGTGCAGATGCTGCTGCTGAAACACAGAGCTCAGCGCTTCCAGGGGATTT-- 1080
Db 3120 TCCTGTCTGGTGCAGATGCTGCTGCTGAAACACAGAGCTCAGCGCTTCCAGGGGATTT-- 3177
Qy 1081 NCCAGCCCCCGGCTCARGAGTGGGAGCAGGCTCGNCAGCAGNAGCNAAGGTTGGGG 1140
Db 3178 GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGGGCTCG----CAGCAAGAGTGGGG 3230
Qy 1141 CAAGCNAAGATGCTCTCCAGGATTTCAACACTGAGCCCTGCCCCAC-----CCTGCTGA 1200
Db 3231 GCAAGCAAGATGCTCTCCAGGATTTCAACACTGAGCCCTGCCCCAC-----CCTGCTGA 3284
Qy 1201 AAACAATNCCGCGAGTGAAGAGACAGAGGAGTGNAGGAGTTNNACCTTGGGAA 1260
Db 3285 AAACAATNCCGCGAGTGAAGAGACAGAGGAGTGNAGGAGTGNAGGAGTNNACCTTGGGAA 3340
Qy 1261 ACAAACAGGAGTCTTTTCTCTGCCCCCTCTCAGTNCAGGTTGGCTGNAACCCGCTTGG 1320
```





Qy	541	TCTCCCTGTCGGATGACAGTGTAGAGAAACCCATCAAAAGGCCTCTCAAAGCTTCGCGGAC	600
Db	2640	TCTCCCTGTCGGATGACAGTGTAGAGAAACCCATCAAAAGGCCTCTCAAAGCTTCGCGGAC	2699
Qy	601	CCCTGAGCTCAGGGGTACACTCCCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTTGA	660
Db	2700	CCCTGAGCTCAGGGGTACACTCCCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTTGA	2759
Qy	661	ACATGGTGTGCGCCGGGGCGGCCACACGACACCCCAAGCTATTTCAATGGTGTGAAG	720
Db	2760	ACATGGTGTGCGCCGGGGCGGCCACACGACACCCCAAGCTATTTCAATGGTGTGAAG	2819
Qy	721	TCCAAATACAGTCTCTTTAAATGGTGAACAACCTGCACATCCGGGAGTTCCACCGGGTCAAAG	780
Db	2820	TCCAAATACAGTCTCTTTAAATGGTGAACAACCTGCACATCCGGGAGTTCCACCGGGTCAAAG	2879
Qy	781	TGGGAGACATCGCCACTTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA	840
Db	2880	TGGGAGACATCGCCACTTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA	2939
Qy	841	CCAAAGCAGGCGACGCTGTTTCGCTACGACATGGAAGTGCACAGCTCGGGCATCGACTGC	900
Db	2940	CCAAAGCAGGCGACGCTGTTTCGCTACGACATGGAAGTGCACAGCTCGGGCATCGACTGC	2999
Qy	901	AGTGCACACTGGCCCTTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCATTGGCCAGC	960
Db	3000	AGTGCACACTGGCCCTTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCATTGGCCAGC	3059
Qy	961	TGGAGAACAGGCCCTTAACCTTGCCCTCCACGGCCGGCTCCACATCTCGCGAAAGCAGCCT	1020
Db	3060	TGGAGAACAGGCCCTTAACCTTGCCCTCCACGGCCGGCTCCACATCTCGCGAAAGCAGCCT	3119
Qy	1021	TCCTGCTTCGGTGCACGATGCTGCCCTCGAAACACAGGCTCAGCCGTTCCACAGGGGATYTG	1080
Db	3120	TCCTGCTTCGGTGCACGATGCTGCCCTCGAAACACAGGCTCAGCCGTTCCACAGGGGATY--	3177
Qy	1081	NCCAGCCCCCGGCTCARGAGTGGGAAACAGGGCCTTCGNCAGCNAGCNAGGTTGGGGG	1140
Db	3178	GCCAGCCCCCGGCTCA--CAGTGGGAAACAGGGCCTCG----CAGCACAAGGTGGGG	3230
Qy	1141	CAAGCNAGATGCTCTCCAGGATTTACANCCTGAGCCNTGCCCCANCCCTGCTGAADA	1200
Db	3231	GCAAGCAGAATGCTCTCCAGGATTTACACTGAGCCCTCGCCAC-----CCTGCTGA	3284
Qy	1201	AAACAYTNCGCCACGTTGAAGACAGAGGAGATGGNCAGGAGTTNNACCTYGGGAA	1260
Db	3285	AAAACATCCGACAGTGAAGAGACAGAGAGATGTC-----AGGAGTTACCTTGGGAA	3340
Qy	1261	ACAAACAGGGATCTTTNTTCTGCCCTGCTCCAGTNCGAGTTGGCTGNACCCGCTTGG	1320
Db	3341	ACAAAACAGGGATCTTTT--TTCTGCCCTGCTCCAGT--CGAGTTGGCCTGA--CCCGCTTG	3396
Qy	1321	ANTCAGTGAACATTTGTTGGCAGANCAGGGGAGCAGCTTCCAGCTCGGGTCAGAAGGG	1380
Db	3397	GATCAGTGAACCATTTGTTGGCAGA--CAGGGGAGCAGCTTCCAGCCTGGGTGAGAAGGG	3455
Qy	1381	GTGGGCGAGCCCTTCGCGCCCTCACCTNCAGCTCTGTGNAGAGTGTCAAGTGTGA	1440
Db	3456	GTGGGCGAGCCCTTCGCGCCCTCACCT--CGAGCTGCTGTG--AGAGTGTCAAGTGTGA	3513
Qy	1441	AGGNCCCAANCTCAGGNTTCAGTGAGAACCAAGGTTNCAGCAGGTATGCCCGCCGNTA	1500
Db	3514	AGGSCCCAAATC--AGGTTTCAGTGCAGAACCAAGGT--CAGCAGGTATGCCCGCCGTAG	3569
Qy	1501	GTTTAAANNNGGGGCCCTCTNAAACCCCTTGGCTNNGGCTTNCACCTNNGCAGCTCANCCC	1560
Db	3570	GTTTAA----GGGGGCCCTCTAAAACCCCTTGCTGCGCTCA--CCTGGCCAGGTCA--CCC	3621
Qy	1561	CTTTTGGGTGTAGGGGAAAAGATGCTGACCTCGGGAAGGCTWCCCTGTTAGATAACAC	1620
Db	3622	CTTTTGGGTGTAGGGGAAAAGATGCTGACCTCGGGAAGGCT--CCCTGGTAGATAACAC	3680
Qy	1621	CACACTTTTTCAGGTGTGTGCAACACAGGTCCTGAGTTTGACCTCTGGTTTCAGCAAGGACC	1680

Db	3681	 CACACATTTTCAGGTGTTGGTGCACACAGAGTCCTGAGTTGCACCTCTGGTTTCACCCAGGAGCC	3740
Qy	1681	 AAAGAAGGTGCTAAAGTGAAGTGGTTCTCAGTNCGCCACAGATGTGCCCTTTGCTGCTG	1740
Db	3741	 AAAGAAGGTGCTAAAGTGAAGTGGTTCTCAGT-CCCCAGACATGTGCCCCCTTGTGCTG	3799
Qy	1741	 GCTACCACTTTTCCCAGACAGCAGGCCCCGAGCCCCCTTCAGGCCACAGCACTGCCCCAG	1800
Db	3800	 GCTACCACTTTTCCCAGACAGCAGGCCCCGAGCCCCCTTCAGGCCACAGCACTGCCCCAG	3859
Qy	1801	 ACTCGCTGGCACTCAGTTCCCTCATCTCTAAAGGTGAAGGTGATGCAGGATATGCCTGA	1860
Db	3860	 ACTCGCTGGCACTCAGTTCCCTCATCTCTAAAGGTGAAGGTGATGCAGGATATGCCTGA	3919
Qy	1861	 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGGNAAGACAGCAGAGAGAGAGY	1920
Db	3920	 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA--GGAAGCAGCAGAGAGAGAGC-	3976
Qy	1921	 TCCGGGCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTCTNGTTNCCCAGNAGACAC	1980
Db	3977	 TCCGGGCCCCAGNCCCACCTATCAGTGT---CCAGCGTGTGGTTTCCCCAGAGCACAGC-	4032
Qy	1981	 GCTNCAGNCACTCAGCACTGACACTNCACTTCGCTCCCTCCCTGCGCCAGCAGGACTCTG	2040
Db	4033	 -----TCAGCATCACACTGACACTCAGCTGCCCTGCCCTGCG--CCAGAGGGTACTG	4083
Qy	2041	 CCGNACGGCACTTTGCACNTCTGATGNAACCTCAAAGCACTTTCATGGCTNGCCCTCTNNG	2100
Db	4084	 CCG-ACGGCACTTTGCACCT-CTGATGACCTCAAAGCACTTTCATGGCTGCCCTCTGGCA	4140
Qy	2101	 GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAAGCCAGAGATGG	2160
Db	4141	 GGCAGGCGGCGGCGAGTGACACT-----GTAGAGCATAGCAAGCCAGAGATGG	4190
Qy	2161	 GGTGAAGGGAACACAGTCTTCAGCTGTCCANCATGCATGTCACTNCCCTCAAACTCTTN	2220
Db	4191	 GGTGAA--GGGACACAGTCTTGAGCTGTCCA-CATGCATGTGACT-CCTCAAACTCTT-	4245
Qy	2221	 NCCAGNATTTCTTAAGAATAGCAGNCCCCTTNCGCCATTTGCCCGAGCTTAGCCTCTTCT	2280
Db	4246	 --CCAGATTTCTTAAGAATAGCACCCCT--TTCCCCATTTGCCCGAGCTTAGCCTCTTCT	4301
Qy	2281	 CCAGGGGAGCTANTCAGGACTCAGCTAGACATTAAATCAGCTGTGNAATCGTCAGGGGG	2340
Db	4302	 CCAGGGGAGCTA-CTCAGGACTCAGCTAGACATTAAATCAGCTGTG-AATCGTCAGGGGG	4359
Qy	2341	 TGCTGCTAGCCTCAACCTCTCTGGGGCAGGGGACCGGAGACTCCGTGGGAGAAAGCTCAT	2400
Db	4360	 TGCTGCTAGCCTCAACCTCTCTGGGGCAGGGGACCGGAGACTCCGTGGGAGAAAGCTCAT	4419
Qy	2401	 TCCCACTCTTGGCAAGACAGGCTTTNGTCCAGCTGTCCAATTGAGTCAGACTGCTCCC	2460
Db	4420	 TCCCACTCTTGGCAAGACAGGCTTTT-GTCCAGCTGTCCAATTGAGTCAGACTGCTCCC	4478
Qy	2461	 GGGGAGAGACCCCGGCCCCAGCACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG	2520
Db	4479	 GGGGAGAGACCCCGGCCCCAGCACATAAAGAACTGCAGCTTGGTACTGCAGAGTCTG	4538
Qy	2521	 GGTTGTAGAAACTCTTTGTAAAGCAATAAAGTTTGGGTGTGATGACAAATGTTAAAAA	2578
Db	4539	 GGTTGTAGAAACTCTTTGTAAAGCAATAAAGTTTGGGTGTGATGACAAATGTTAAAAA	4596

```

RESULT 8
US-09-053-375B-1234
; Sequence 1234, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chencik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B

```

```

; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Ver
; SEQ ID NO 1234
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-053-375B-1234

```

Query Match 70.3%; Score 1850.2; DB 16; Length 4596;

Qy	1	CCCCCTCTCAGAGCCCGGCGCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	60
Db	2100		2159
		CCCCCTCTCAGAGCCCGGCGCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	
Qy	61	CTCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAGTGGGAGGTCTGAAGA	120
Db	2160		2219
		CTCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAGTGGGAGGTCTGAAGA	
Qy	121	GCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGAGCAAATTACC	180
Db	2220		2279
		GCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGAGCAAATTACC	
Qy	181	ACCAGACCTCTCCATGCCCGAGCGAGAGAGCTTTGCGCAAGGGCCCCCAGGGCCCGGCCAG	240
Db	2280		2339
		ACCAGACCTCTCCATGCCCGAGCGAGAGAGCTTTGCGCAAGGGCCCCCAGGGCCCGGCCAG	
Qy	241	CTGAGGAGACAAACAGGCAGAGAGCCCTTAAGCTCAGACCTCTCTCCACACAGAGAGCCCCCAG	300
Db	2340		2399
		CTGAGGAGACAAACAGGCAGAGAGCCCTTAAGCTCAGACCTCTCTCCACACAGAGAGCCCCCAG	
Qy	301	AGCCAAACAAGTCTCCCTTGACTTTGAGCAAGAGGAGGTCTGGGATCTGGGAAACCCCT	360
Db	2400		2459
		AGCCAAACAAGTCTCCCTTGACTTTGAGCAAGAGGAGGTCTGGGATCTGGGAAACCCCT	
Qy	361	TACCTCTGTCTCCTCGAGCCAGCCCTGCCAGAAACCCACAGCTCACACAGAGCGGAAAG	420
Db	2460		2519
		TACCTCTGTCTCCTCGAGCCAGCCCTGCCAGAAACCCACAGCTCACACAGAGCGGAAAG	
Qy	421	CAACCGTCCCGGAGCAGGAACTGCGAGCAGCTGGAAATAGAAATTATTCCTCAACAGACCTGT	480
Db	2520		2579
		CAACCGTCCCGGAGCAGGAACTGCGAGCAGCTGGAAATAGAAATTATTCCTCAACAGACCTGT	
Qy	481	CCAGACCATTTCTCTGGAGGACAGAGCAAAATCTCTCGTCCTCAGCATCGACAGCC	540
Db	2580		2639
		CCAGACCATTTCTCTGGAGGACAGAGCAAAATCTCTCGTCCTCAGCATCGACAGCC	
Qy	541	TCTCCCTGTCCGATGACAGTGAGAAGAAACCATCAAAGGCTCTCAAAGCTCCGGGACA	600
Db	2640		2699
		TCTCCCTGTCCGATGACAGTGAGAAGAAACCATCAAAGGCTCTCAAAGCTCCGGGACA	
Qy	601	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCGCCGAGGCTCGAAGCTCCAGCTGGGA	660
Db	2700		2759
		CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCGCCGAGGCTCGAAGCTCCAGCTGGGA	
Qy	661	ACATGTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	720
Db	2760		2819
		ACATGTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	
Qy	721	TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCACCGGGTCAAAG	780
Db	2820		2879
		TCCAAATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCACCGGGTCAAAG	
Qy	781	TGGGAGACATCGCCACTGGGCATCAGCAGCCAGATCCCAGCTGCAGCCTTCAGCTTGGTCA	840
Db	2880		2939
		TGGGAGACATCGCCACTGGGCATCAGCAGCCAGATCCCAGCTGCAGCCTTCAGCTTGGTCA	
Qy	841	CCAAAGACGGGAGCCTGTTGCTAGCAATGAGAGGTGCCAGACTCGGGCATCGACTCTGC	900
Db	2940		2999
		CCAAAGACGGGAGCCTGTTGCTAGCAATGAGAGGTGCCAGACTCGGGCATCGACTCTGC	

```

4033 -----TCAGATCACACTGACACTCACCTGCGCTGCCCTGCTG--CCAGAGGGTACTG 4083
2041 CCGNACGGACATTGTCANVCTGATGNACCTCAAGACACTTTCATGCTGCTGCTCTNNNG 2100
4084 CCG-ACGGCACTTTGCACT--CTGATGACCTCAAGACACTTTCATGCTGCGCTCTGGCA 4140
2101 GCAGGNCAGGNCAGGNCAGTGCACNCTGTAGNAGCATANGACAANGCCAGGAGATGG 2160
4141 GGGCAGGGCAGGGCAGTGCACCT-----GTAGAGCATAGCAAGCCAGGAGATGG 4190
2161 GGTGNAAGGANCACAGCTTTGAGCTTGCANCATGATGACTNCTCAAACTCTTN 2220
4191 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGCAATGTGACT-CCTCAAACTCTT- 4245
2221 NCCAGNATTTCTTAAGATAGCANCCTTNCCTTCCCTGCTGCTGCTGCTTCT 2280
4246 --CCAGATTTCTTAAGATAGCACCCTT--TTCCCAATTGCCCCAGCTTAGCTTCTTCT 4301
2281 CCCAGGGAGCTANCTCAGGACTCAGTAGCATTAAATCAGCTGTGNAATCGTCAGGGGG 2340
4302 CCCAGGGAGCTA-CTCAGGACTCAGTAGCATTAAATCAGCTGTG-AACTGTCAGGGGG 4359
2341 TGTCTGCTAGCTCAACCTCTTGGGCGAGGAGCGCGAGACTCCGTCGAGAGCTCAT 2400
4360 TGTCTGCTAGCTCAACCTCTTGGGCGAGGAGCGCGAGACTCCGTCGAGAGCTCAT 4419
2401 TCCACATCTTGCAAGACAGCTTTNGTCCAGCTGTCCACATTCAGTGCAGCTGCTCCC 2460
4420 TCCACATCTTGCAAGACAGCTTT-GTCCAGCTGTCCACATTCAGTGCAGCTGCTCCC 4478
2461 GGGGAGAGAGCCCGGCCCGCCAGCACATAAAGAACTGAGCTGCTGCTGCTGCTGCTG 2520
4479 GGGGAGAGAGCCCGGCCCGCCAGCACATAAAGAACTGAGCTGCTGCTGCTGCTGCTG 4538
2521 GGTGTAGAGAACTCTTGTAAAGTAAAGTTGGGGTGATGACAAATGTTAAAAA 2578
4539 GGTGTAGAGAACTCTTGTAAAGTAAAGTTGGGGTGATGACAAATGTTAAAAA 4596

RESULT 9
US-09-949-003C-437
; Sequence 437, Application US/09949003C
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH NEUROPSYCHIATRIC
; DISORDERS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000791
; CURRENT APPLICATION NUMBER: US/09/949,003C
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,446
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 74065
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 437
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-949-003C-437

Query Match 70.3%; Score 1850.2; DB 41; Length 4596;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY 1 CCCCTCTCACAGCCCGAGGCTTCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCTGT 60
DB 2100 CCCCTCTCACAGCCCGAGGCTTCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCTGT 2159
QY 61 CTGCAAGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120
DB 2160 CTGCAAGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2219
QY 121 GCCCTTGGAGGGGAGATATATAAGAAACCAAGACATCCACCCCAATCAAGCCCAATTACC 180
```

```

DB 2220 GCCCTTGGAGGGGAGATATATAAGAAACCAAGACATCCACCGCAATCAAGCCCAATTACC 2279
QY 181 ACCAGACCTTCATGCTCCAGCCGAGAGAGCTTTTCCCAAGGGCCCGAGGGCCCGGCCAG 240
DB 2280 ACCAGACCTTCATGCTCCAGCCGAGAGAGCTTTTCCCAAGGGCCCGAGGGCCCGGCCAG 2339
QY 241 CTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCCGCCAG 300
DB 2340 CTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCCGCCAG 2399
QY 301 AGCCAAACAAGTCTCTCTCTCTTGTAGCAAGAGAGGAGTCTGGGATGTGGGAACCTT 360
DB 2400 AGCCAAACAAGTCTCTCTCTCTTGTAGCAAGAGAGGAGTCTGGGATGTGGGAACCTT 2459
QY 361 TACCTCTGCTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAG 420
DB 2460 TACCTCTGCTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAG 2519
QY 421 CAACCGTCCCGAGAGCAGGAACCTGACAGCTGGAATAGAAATATTCTCAACAGCCTGT 480
DB 2520 CAACCGTCCCGAGAGCAGGAACCTGACAGCTGGAATAGAAATATTCTCAACAGCCTGT 2579
QY 481 CCCAGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGCTGCCCTCAGCATCGACGCC 540
DB 2580 CCCAGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGCTGCCCTCAGCATCGACGCC 2639
QY 541 TCTCCCTGTCCGAGTGCAGTGAGAAACCCATCAAGGGCTCTCAAAGCTCCCGGGACA 600
DB 2640 TCTCCCTGTCCGAGTGCAGTGAGAAACCCATCAAGGGCTCTCAAAGCTCCCGGGACA 2699
QY 601 CCCTGAGCTCAGGCGTACACTCTCTGAGAGCAGCAGCGAGGCTCGAAGCTCCAGCTGGA 660
DB 2700 CCCTGAGCTCAGGCGTACACTCTCTGAGAGCAGCAGCGAGGCTCGAAGCTCCAGCTGGA 2759
QY 661 ACATGCTGCTGGCCCGGGGGGCCACCCAGCTATTTCAATGCTGTGAAG 720
DB 2760 ACATGCTGCTGGCCCGGGGGGCCACCCAGCTATTTCAATGCTGTGAAG 2819
QY 721 TCCAAATACAGTCTCTTAATGGTGAACACCTTGCACTCCGGGAGTTCCACCGGGTCAAAG 780
DB 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTTGCACTCCGGGAGTTCCACCGGGTCAAAG 2879
QY 781 TGGGAGACATCGCACACTGGGCATCAGAGCCAGATCCAGCTGCAAGCTTCAGCTTGTGCA 840
DB 2880 TGGGAGACATCGCACACTGGGCATCAGAGCCAGATCCAGCTGCAAGCTTCAGCTTGTGCA 2939
QY 841 CCAAGACGGCAGCGCTGTTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900
DB 2940 CCAAGACGGCAGCGCTGTTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 2999
QY 901 AGTGACACACTGGCCCTGATGGCAGCTTCGCTTGGAGCTGGAGGTCAAGCATGCCAGC 960
DB 3000 AGTGACACACTGGCCCTGATGGCAGCTTCGCTTGGAGCTGGAGGTCAAGCATGCCAGC 3059
QY 961 TGGAGAACAGGCCCTAACCTGCCCTCCACCGCGGCTCCACACTGCCCGGAAAGCAGCT 1020
DB 3060 TGGAGAACAGGCCCTAACCTGCCCTCCACCGCGGCTCCACACTGCCCGGAAAGCAGCT 3119
QY 1021 TCCTGCTCGGTGCACGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG 1080
DB 3120 TCCTGCTCGGTGCACGATGCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATYTG 3177
QY 1081 NCCAGCCCCCGGCTCARCAGNTGGGAACCCAGGGCTTCGNCAGCNAGCNAAAGTTGGGGG 1140
DB 3178 GCCAGCCCCCGGCTCA--CAGTGGGAACCCAGGGCTCG-----CAGCAGCAAGTGGGG 3230
QY 1141 CAAGCNAGAAATGCTCCCAAGGATTTCAACNCTGAGCCNCTGCCCCANCCCTGCTGAADA 1200
DB 3231 GCNAGCAAGATGCTCCCAAGGATTTCAACCTGAGCCCTGCCCCAC-----CCTGCTGA 3284
QY 1201 AAACATNCCGCCCATGTAAGAGACAGAGGAGGTGNCAGGAGTTNNACCTYGGGGAA 1260
```

Db 3285 AAAAAATCCCGACCTGAGAGACAGAGGAGATGGC-----AGGAGTTACCTGGGGAA 3340  
Qy 1261 AAAAAACAGGAGATCTTNTTCTGCCCCCTGCTCCAGTNCGAGTTGGCCCTGNAACCGCTTGG 1320  
Db 3341 AAAAAACAGGAGATCTTT-TTCTGCCCCCTGCTCCAGT--CGAGTTGGCCTGA--CCCGCTTG 3396  
Qy 1321 ANTCAAGTACCATTTGTTGGCAGANAGGAGAGAGCAGCTTCCAGCCCTGGGTCAAGAGG 1380  
Db 3397 GATCAGTGAACCATTTGTTGGCAGA--CAGGGAGAGCAGCTTCCAGCCCTGGGTCAAGAGG 3455  
Qy 1381 GTGGGCGAGCCCTTCGGGCCCTCTACCCCTNCCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440  
Db 3456 GTGGGCGAGCCCTTCGGGCCCTCTACCCCT--CCAGGCTGCTGTG--AGAGTGTCAAGTGTGA 3513  
Qy 1441 AGGNCCTCAANTCAGNTTTCAGTGCAGAACCAAGTNCAGCAGTATGCCCCCGCNGTA 1500  
Db 3514 AGGCCCAAACTC---AGGTTCAAGTGCAGAACCAAGT--CAGCAGGTATGCCCCCGCTAG 3569  
Qy 1501 GGTAAANNNGGGGCCCTCTNAAACCCCTTGCTNGGCTNCACTNGGCCAGCTCANCCC 1560  
Db 3570 GTTAA-----GGGGGCCCTCTAAACCCCTTGCTGGCCCTCA---CCTGGCCAGCTCA--CCC 3621  
Qy 1561 CTTTGTGGGTGAGGGGAAAAGAAATGCCTGACCCCTGGGAAGGCTWCCCTGTGTAGAAATACAC 1620  
Db 3622 CTTTGTGGGTGAGGGGAAAAGAAATGCCTGACCCCTGGGAAGGCT--CCCTGTGTAGAAATACAC 3680  
Qy 1621 CACACTTTTTCAGGTTGTTGCAACACAGGTCTGTGAGTTGACCTCTGTGTTTCAGCCAAAGACC 1680  
Db 3681 CACACTTTTTCAGGTTGTTGCAACACAGGTCTGTGAGTTGACCTCTGTGTTTCAGCCAAAGACC 3740  
Qy 1681 AAAGAAGGTGTGAAGTGAAGTGTCTTCAGTNCCTCCAGACATGTGCCCCCTTTGCTGCTG 1740  
Db 3741 AAAGAAGGTGTGAAGTGAAGTGTCTTCAGT--CCCCAGACATGTGCCCCCTTTGCTGCTG 3799  
Qy 1741 GCTACCACTCTTCCCCAGAGCAGAGCCCGAGCCCTTTCAGGCCAGCACTGCCCCCAG 1800  
Db 3800 GCTACCACTCTTCCCCAGAGCAGAGCCCGAGCCCTTTCAGGCCAGCACTGCCCCCAG 3859  
Qy 1801 ACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGGTGAAGGTGTATGAGGATATGCTCTGA 1860  
Db 3860 ACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGGTGAAGGTGTATGAGGATATGCTCTGA 3919  
Qy 1861 CAGAACAGTCTGTGGATGACATGATCAGTGTCTNAAAGNAAAGCAGCAGAGAGACGY 1920  
Db 3920 CAGAACAGTCTGTGGATGACATGATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGACG-- 3976  
Qy 1921 TCCGGGCCCCAGNCCCACCTNATCAGTGTNCCAGCGTGTCTGTTNCCCGAGNAGACA 1980  
Db 3977 TCCGGGCCCCAGNCCCACCTATCAGTGT---CCAGCGTGTGTTTCCCGAGAGCACAGC-- 4032  
Qy 1981 GCTNCAAGCATCANCACTGACATNCACTTNGCCCTGCCCCCTNGGCCANGAGGGTACTG 2040  
Db 4033 -----TCAGCATCACTGACACTCACCTGCCCCCTGCCCCCTG--CCAGAGGGTACTG 4083  
Qy 2041 CCGNACGGCATTTTGCACNTCTGATGNACTCTCAAGCACTTTTCATGCTNGCCCTCTNNG 2100  
Db 4084 CCG--ACGGCATTTTGCACNTCTGATGNACTCTCAAGCACTTTTCATGCTNGCCCTCTGCA 4140  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCACANCTGTAGGNCAGCATNCCANGCCAGGAGATGG 2160  
Db 4141 GGGCAGGGCAGGGCAGTGAACACT-----GTAGGAGCATAGCAAGCCAGGAGATGG 4190  
Qy 2161 GGTGNAAGGGANACAGCTCTTGTAGCTGTCCANCATGATGTGACTNCTCAAACTCTTN 2220  
Db 4191 GGTGAA--GGGACACAGCTTTGTAGCTGTCCA--CATGATGTGACT--CTCAAACTCTT-- 4245  
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCCCCTTNGCCCATTTGCCCCAGCTTTAGCTCTTCT 2280  
Db 4246 --CCAGATTTCTTAAGATAGCACCCCC--TTCCCCCATTTGCCCCAGCTTTAGCTCTTCT 4301  
Qy 2281 CCGAGGGAGCTANTCTCAGACTCAGCTAGCATTAATCAGCTGTGNAATCGTCAGGGG 2340  
Db 4302 CCGAGGGAGCTA--CTCAGGACTCAGCTAGCATTAATCAGCTGTG--AATCGTCAGGGG 4359

Qy 2341 TGCTGTAGCTCAACCTCTCTGGGCGAGGGGACGCCGAGACTCCCTGTGGGAGAGCTCAT 2400  
Db 4360 TGCTGTAGCTCAACCTCTCTGGGCGAGGGGACGCCGAGACTCCCTGTGGGAGAGCTCAT 4419  
Qy 2401 TCCACATCTTGGCAGAGAGCCCTTNGTCCAGCTGTCCACATTTGAGTGCAGCTGCTCCC 2460  
Db 4420 TCCACATCTTGGCAGAGAGCCCTT--GTCCAGCTGTCCACATTTGAGTGCAGCTGCTCCC 4478  
Qy 2461 GGGGAGAGAGCCCGGCCCCAGCACATAAAGAACTGACGCTTGGTACTGCAGAGTCTG 2520  
Db 4479 GGGGAGAGAGCCCGGCCCCAGCACATAAAGAACTGACGCTTGGTACTGCAGAGTCTG 4538  
Qy 2521 GGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 2578  
Db 4539 GGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 4596

## RESULT 10

US-09-981-397A-17  
; Sequence 17, Application US/09981397A  
; GENERAL INFORMATION:  
; APPLICANT: Axxima Pharmaceuticals AG  
; APPLICANT: Schubart, Daniel  
; APPLICANT: Habenberger, Peter  
; APPLICANT: Stein-Gerlach, Matthias  
; APPLICANT: Bevec, Dorian  
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their  
; TITLE OF INVENTION: Inhibition  
; FILE REFERENCE: AXM-004.1 US  
; CURRENT APPLICATION NUMBER: US/09/981,397A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/240,750  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-981-397A-17

Query Match 70.3%; Score 1850.2; DB 45; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCTGT 60  
Db 2100 CCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCTGT 2159  
Qy 61 CTGACGCGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTCTGAAGA 120  
Db 2160 CTGACGCGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTCTGAAGA 2219  
Qy 121 GCCTTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATACC 180  
Db 2220 GCCTTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATACC 2279  
Qy 181 ACCAGACCTTCCATGCCCGAGAGAGCTTTGCGCCAGGGCCCGAGGGCCCGGCGAG 240  
Db 2280 ACCAGACCTTCCATGCCCGAGAGAGCTTTGCGCCAGGGCCCGAGGGCCCGGCGAG 2339  
Qy 241 CTGAGGAGACAAAGCAGCAGAGCCCTTAAGTCCAGGCTCTCTCCACACAGAGCCCCCAG 300  
Db 2340 CTGAGGAGACAAAGCAGCAGAGCCCTTAAGTCCAGGCTCTCTCCACACAGAGCCCCCAG 2399  
Qy 301 AGCCAAACAGTCTCTCTCTTGAAGCAAGGAGTCTGGGATGTGGAAACCT 360  
Db 2400 AGCCAAACAGTCTCTCTCTTGAAGCAAGGAGTCTGGGATGTGGAAACCT 2459  
Qy 361 TACCTCTGCTCTCTCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACCAGAGCGGAAG 420  
Db 2460 TACCTCTGCTCTCTCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACCAGAGCGGAAG 2519

Qy	421	CAACCGTCCCGAGCAGGAAC	CTCAGCAGCTGGAAATAGAA	TATTTCCTCAACAGCCTGT	480
Db	2520	CAACCGTCCCGAGCAGGAAC	CTCAGCAGCTGGAAATAGAA	TATTTCCTCAACAGCCTGT	2579
Qy	481	CCAGACCAATTTTCTCTGGAGGAGCAGAG	CAAAATTTCTCTCGTGCCTCAGCATCGACAGCC	540	
Db	2580	CCAGACCAATTTTCTCTGGAGGAGCAGAG	CAAAATTTCTCTCGTGCCTCAGCATCGACAGCC	2639	
Qy	541	TCCTCCCTGTCGGATGACAGTGAGAAGAAC	CCATCAAAAGCCTCTCAAAGCTCGCGGACA	600	
Db	2640	TCCTCCCTGTCGGATGACAGTGAGAAGAAC	CCATCAAAAGCCTCTCAAAGCTCGCGGACA	2699	
Qy	601	CCCTTGAGCTCAGCGCTACACTCTCTGGAGCAGCGCCAGGCTCGAAGCTCCAGTGA	660		
Db	2700	CCCTTGAGCTCAGCGCTACACTCTCTGGAGCAGCGCGGAGGCTCGAAGCTCCAGTGA	2759		
Qy	661	ACATGTGTGGCCCGGGGGCGGCCACCGACAC	CCCAAGCTATTTCAATGTGTGAAG	720	
Db	2760	ACATGTGTGGCCCGGGGGCGGCCACCGACAC	CCCAAGCTATTTCAATGTGTGAAG	2819	
Qy	721	TCCAAATACAGTCTCTTAATGTGTAAACAC	CTGCACATCCGGGAGTTTCCACCGGCTCAAAAG	780	
Db	2820	TCCAAATACAGTCTCTTAATGTGTAAACAC	CTGCACATCCGGGAGTTTCCACCGGCTCAAAAG	2879	
Qy	781	TGGGAGACATCGCCACTGGCATCAGCAGCGAGAT	CCGAGCTGCAGCGCTTGAAGCTTGCTCA	840	
Db	2880	TGGGAGACATCGCCACTGGCATCAGCAGCGAGAT	CCGAGCTGCAGCGCTTGAAGCTTGCTCA	2939	
Qy	841	CCAAAGACGGGACGCTGTTTCGTACGACATGAGGTGC	CAGACTCGGCGATCGACCTGC	900	
Db	2940	CCAAAGACGGGACGCTGTTTCGTACGACATGAGGTGC	CAGACTCGGCGATCGACCTGC	2999	
Qy	901	AGTGCAACATGGCCCTGTATGGCAGCTTCCGCTTGGAGCT	GGAGGCTCAAGCATGGCCAGC	960	
Db	3000	AGTGCAACATGGCCCTGTATGGCAGCTTCCGCTTGGAGCT	GGAGGCTCAAGCATGGCCAGC	3059	
Qy	961	TGGAGAACAGGCGCTTAACCTTGCCCTTCCA	CGCGCGGCTCCACCTGCGGAAAGCAGCT	1020	
Db	3060	TGGAGAACAGGCGCTTAACCTTGCCCTTCCA	CGCGCGGCTTCCACCTGCGGAAAGCAGCT	3119	
Qy	1021	TCCTGCTCGGTGCACGATGCTGCCCTGAAACAC	AGGCTCAGCGCTTCCAGGGGATYTG	1080	
Db	3120	TCCTGCTCGGTGCACGATGCTGCCCTGAAACAC	AGGCTCAGCGCTTCCAGGGGATYTG	3177	
Qy	1081	NCCAGCCCCGGCTCARGAGTGGGAA	CAGGCGCTTCGNACGNAGGATNGGGG	1140	
Db	3178	GCCAGCCCCCGGCTCA	- CAGTGGGAAACAGGCGCTTCG - - - - CAGCAGCAAGGTGGG	3230	
Qy	1141	CAAGCNAGATGCTCTCCAGGATTTCA	CANCCTGAGCCNTGCCCCANCCCTCTGAADA	1200	
Db	3231	GCAAGCAGATGCTCTCCAGGATTTCA	CAGCTGAGCCCTCGCCAC - - - - - CCTGCTGA	3284	
Qy	1201	AAACAYTNCGCCACGTGAAGAGACAGAGGAT	TGNCAGGAGTTNNACCTYGGGAA	1260	
Db	3285	AAAAACATCCCGACGTGAAGAGACAGAGGATGCG	- - - - - AGGAGTTACCTGGGAA	3340	
Qy	1261	ACAAAACAGGGATCTTTNTCTGCCCTTGCTCCA	GTNCAGTGTGGCTGNAACCGCTTGG	1320	
Db	3341	ACAAAACAGGGATCTTTNTCTGCCCTTGCTCCA	GTNCAGTGTGGCTGNAACCGCTTGG	3396	
Qy	1321	ANTCAGTGACCATTTGTTGGCAGAN	CAGGGAGACGCTTCCAGCTGGGTGAGAGGG	1380	
Db	3397	GATCAGTGACCATTTGTTGGCAGAN	CAGGGAGACGCTTCCAGCTGGGTGAGAGGG	3455	
Qy	1381	GTGGCGAGCCCTTCGGCCCCCTCACCCTTCC	AGGCTGCTGTGNAGAGTGTCAAGTGTGTA	1440	
Db	3456	GTGGCGAGCCCTTCGGCCCCCTCACCCTTCC	AGGCTGCTGTGNAGAGTGTCAAGTGTGTA	3513	
Qy	1441	AGGNVCCAAANCTCAGGNTTCAGTGACAGAA	CCAGGTCAGAGGATGCGCCCGGNTA	1500	
Db	3514	AGGGCCAAAACCTC - - - AGGTTCA	GTGAGAACCAAGT - CAGCAGGATGCGCCCGGTAG	3569	

QY	1501	GGTTAA	NNGGGGG	CCCTCT	TAAACCC	CTTGCC	TNGCC	TNCAC	TNGG	CCAG	CTCAN	CCC	1560	
DB	3570	GT	TAA----	GGGG	CCCT	TAAACCC	CTTGCC	TNGCC	TCA--	CC	TGGG	CAG	CTCA-CCC	3621
QY	1561	CTTTTGG	TGTAG	GGGAAA	GAATG	CTGAC	CTGG	GAAG	GGCTT	WC	CGT	TGG	CAAGG	1610
DB	3622	CTTTTGG	TGTAG	GGGAAA	GAATG	CTGAC	CTGG	GAAG	GGCTT	CC	CT	TGG	TAG	1620
QY	1621	CAC	CTTTT	CAG	GTGTG	CAACAC	AG	GTCT	GAG	TTG	ACCT	CT	TG	1680
DB	3681	CAC	CTTTT	CAG	GTGTG	CAACAC	AG	GTCT	GAG	TTG	ACCT	CT	TG	3740
QY	1681	AA	GAA	GGTGT	TAAG	TGA	GTGG	TCT	CAG	TN	CCCC	GAC	ATG	1740
DB	3741	AA	GAA	GGTGT	TAAG	TGA	GTGG	TCT	CAG	TN	CCCC	GAC	ATG	3799
QY	1741	GCT	ACCA	CTCT	TCC	CAG	CAG	CAG	CCCG	CAG	CCCTT	CAG	CCCG	1800
DB	3800	GCT	ACCA	CTCT	TCC	CAG	CAG	CAG	CCCG	CAG	CCCTT	CAG	CCCG	3859
QY	1801	ACT	CG	TGG	CACT	CAG	TTCC	TCT	CACT	GT	TA	AG	GGTGA	1860
DB	3860	ACT	CG	TGG	CACT	CAG	TTCC	TCT	CACT	GT	TA	AG	GGTGA	3919
QY	1861	CAG	AA	CAG	CTCT	GT	GAT	GA	CAT	GAT	CA	GT	GT	1920
DB	3920	CAG	AA	CAG	CTCT	GT	GAT	GA	CAT	GAT	CA	GT	GT	3976
QY	1921	TCC	GG	CG	CCC	CAG	NCCC	CA	TN	AT	CA	GT	GT	1980
DB	3977	TCC	GG	CG	CCC	CAG	NCCC	CA	TN	AT	CA	GT	GT	4032
QY	1981	GCT	NC	AG	N	CAT	C	AT	G	CA	TN	C	AC	2040
DB	4033	-----	TC	AG	CA	TC	AC	TC	CA	CT	CA	CT	CA	4083
QY	2041	CCG	NA	CG	CA	CTTT	TG	CA	NC	TT	CA	AG	CA	2100
DB	4084	CCG	NA	CG	CA	CTTT	TG	CA	NC	TT	CA	AG	CA	4140
QY	2101	GC	AG	GC	AG	GC	AG	GC	AG	GC	AG	GC	AG	2160
DB	4141	GGG	CAG	GGC	AG	GGC	AG	GGC	AG	GGC	AG	GGC	AG	4190
QY	2161	GGT	NA	AG	GG	AN	CA	CA	GT	CT	TC	CA	NC	2220
DB	4191	GGT	NA	AG	GG	AN	CA	CA	GT	CT	TC	CA	NC	4245
QY	2221	NCC	AG	NA	TTT	CT	TA	GA	AT	AG	CA	NC	CC	2280
DB	4246	--	CC	AG	AT	TT	CT	TA	GA	AT	AG	CA	NC	4301
QY	2281	CCC	AG	GG	AG	CT	AN	CT	CAG	CA	CT	TA	GA	2340
DB	4302	CCC	AG	GG	AG	CT	AN	CT	CAG	CA	CT	TA	GA	4359
QY	2341	TG	CT	GT	AG	CC	CT	CA	AC	CT	CC	TC	GG	2400
DB	4360	TG	CT	GT	AG	CC	CT	CA	AC	CT	CC	TC	GG	4419
QY	2401	TCC	CA	CAT	CT	TG	CC	AG	CAC	CC	TT	TNG	T	2460
DB	4420	TCC	CA	CAT	CT	TG	CC	AG	CAC	CC	TT	TNG	T	4478
QY	2461	GGG	G	AG	G	AG	CC	CG	CC	CC	CAG	CA	CA	2520
DB	4479	GGG	G	AG	G	AG	CC	CG	CC	CC	CAG	CA	CA	4538
QY	2521	G	TT	T	TA	G	AG	CA	TA	AA	G	TT	T	2578
DB	4539	G	TT	T	TA	G	AG	CA	TA	AA	G	TT	T	4596



## RESULT 11

US-10-172-118-975  
; Sequence 975, Application US/10172118  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veer, Laura  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172,118  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 975  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: NW\_003954  
; DATABASE ENTRY DATE: 2001-06-18  
US-10-172-118-975

Query Match 70.3%; Score 1850.2; DB 49; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;  
  
Qy 1 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60  
Db 2100 CCCCTCTCACAGCCAGCCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2159  
  
Qy 61 CTGACGCGAGCTGGGAGGGAAGTGAAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
Db 2160 CTGACGCGAGCTGGGAGGGAAGTGAAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2219  
  
Qy 121 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCCCAATCAAGCCCAATTACC 180  
Db 2220 GCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCCCAATCAAGCCCAATTACC 2279  
  
Qy 181 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGCCCCAGGGGCCCGGCCAG 240  
Db 2280 ACCAGACCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGGCCCGGCCAG 2339  
  
Qy 241 CTGAGGAGACAAACAGGCAAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCAG 300  
Db 2340 CTGAGGAGACAAACAGGCAAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCCAG 2399  
  
Qy 301 AGCCAAACAACTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGFGGAAACCTT 360  
Db 2400 AGCCAAACAACTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGGATGFGGAAACCTT 2459  
  
Qy 361 TACCTCTCTCTCTCTGAGCAGCCCTGCGCAGAAACCCAGCTCAACAGAGCGGAAAG 420  
Db 2460 TACCTCTCTCTCTCTGAGCAGCCCTGCGCAGAAACCCAGCTCAACAGAGCGGAAAG 2519  
  
Qy 421 CAACCGTCCCGAGCAGGAACTGACGAGCTGGGAAATAGAAATTATCTCAACAGCCTGT 480  
Db 2520 CAACCGTCCCGAGCAGGAACTGACGAGCTGGGAAATAGAAATTATCTCAACAGCCTGT 2579  
  
Qy 481 CCCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTCTGAGCTGAGCAAGCC 540  
Db 2580 CCCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTCTGAGCTGAGCAAGCC 2639  
  
Qy 541 TCTCCCTGTCGATGACAGTACAGAGAACCCATCAAGGGCTCTCAAGCTCGCGGACA 600  
Db 2640 TCTCCCTGTCGATGACAGTACAGAGAACCCATCAAGGGCTCTCAAGCTCGCGGACA 2699  
  
Qy 601 CCCTGAGCTCAGGGGTACACTCTCTGGAGCAGCCAGGGCTCGAAGCTCCAGCTGGA 660

Db 2700 CCCTGAGCTCAGGGGTACACTCTTGAGCAGCCAGGCCAGGCTCGAAGCTCCAGCTGGA 2759  
Qy 661 ACATGGTGTGCGCCCGGGGGGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAAG 720  
Db 2760 ACATGGTGTGCGCCCGGGGGGGCCACCCGACACCCCAAGCTATTTCATGGTGTGAAAG 2819  
Qy 721 TCCAAATACAGTCTCTTAATGGTGAACACCTGSCACATCCGGGAGTTCACCGGGTCAAAG 780  
Db 2820 TCCAAATACAGTCTCTTAATGGTGAACACCTGSCACATCCGGGAGTTCACCGGGTCAAAG 2879  
Qy 781 TGGGAGACATCGCCACTGGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA 840  
Db 2880 TGGGAGACATCGCCACTGGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA 2939  
Qy 841 CCAAAGACGGGCGAGCTGTTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 900  
Db 2940 CCAAAGACGGGCGAGCTGTTCTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC 2999  
Qy 901 AGTGCACACTGGCCCCCTGATGGCAGCTTCGCTTGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3000 AGTGCACACTGGCCCCCTGATGGCAGCTTCGCTTGAGCTGGAGGTCAAGCATGGCCAGC 3059  
Qy 961 TGGAGAACAGGCCCTTAACCTGCCCCCTCCACCGCGGCTCCACACTGCGCGGAAAGCAGCT 1020  
Db 3060 TGGAGAACAGGCCCTTAACCTGCCCCCTCCACCGCGGCTCCACACTGCGCGGAAAGCAGCT 3119  
Qy 1021 TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCGCGCTTCCAGGGGATTTG 1080  
Db 3120 TCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCGCGCTTCCAGGGGATTTG 3177  
Qy 1081 NCCAGCCCCCGGCTCARGAGTGGGAAACAGGGCTCCGNCAGCAGNAGCNAAGGTTNGGGG 1140  
Db 3178 GCACAGCCCCCGGCTCA--CAGTGGGAAACAGGGCTCG-----CAGCAGCAAGTGGGG 3230  
Qy 1141 CAAGCNAAGATGCTCCAGGATTTTCAACCTGAGCCCTGCCCCCAGCTGCTGAADA 1200  
Db 3231 GCAAGCAGAGATGCTCCAGGATTTTCAACCTGAGCCCTGCCCCCAGCTGCTGA 3284  
Qy 1201 AAACAYTNCGCCACGTGAAAGAGACAGAGGATGNCAGAGGATTTNNACTTGGGGAA 1260  
Db 3285 AAAACAYTNCGCCACGTGAAAGAGACAGAGGATGNCAGAGGATTTNNACTTGGGGAA 3340  
Qy 1261 ACAAAACAGGGATCTTTTTCCTGCTGCTCCAGTTCGAGTTCGCTGNAACCTGCTGG 1320  
Db 3341 ACAAAACAGGGATCTTTTTCCTGCTGCTGCTCCAGT--CGAGTTCGCTG--CCGCTTG 3396  
Qy 1321 ANTACGTGACCATTTGTCAGCAGAGGAGAGAGCTTCACAGCTTGGGTCAAGAGGG 1380  
Db 3397 GATCAGTGAACCATTTGTTGGCAGA-CAGGGAGAGCAGCTTCCAGCTTGGGTCAAGAGGG 3455  
Qy 1381 GTGGGCGAGCCCTTCGGGCCCTCACCTTCAGGCTGCTGTGAGAGTGTCAAGTGTGTA 1440  
Db 3456 GTGGGCGAGCCCTTCGGGCCCTCACCTTCAGGCTGCTGTGAGAGTGTCAAGTGTGTA 3513  
Qy 1441 AGGNCACCAANCTCAGGNTTTCAGTGCAAGCAGCTNAGCAGGATGCCCCCGGNTA 1500  
Db 3514 AGGNCACCAANCTC---AGGTTTCAGTGCAAGAACAGGT-CAGCAGGATGCCCCCGGNTAG 3569  
Qy 1501 GGTTAANNNGGGGCGCTCTNAAACCCCTTTCCTNGGCTTNCACCTNGGCCAGCTCANCCC 1560  
Db 3570 GTTAA----GGGGGCGCTTAAACCCCTTTCCTGGCCCTCA--CCTGGCAGCTCA-CCC 3621  
Qy 1561 CTTTTGGGTGAGGGGAAAAGATGCTGACCTTGGGAGGCTGCCCTGGTAGAATACAC 1620  
Db 3622 CTTTTGGGTGAGGGGAAAAGATGCTGACCTTGGGAGGCT--CCCTGGTAGAATACAC 3680  
Qy 1621 CACATTTTTCAGGTGTTGCAACACAGGTCTGAGTGTGAGCTCTGGTTCAGCCAGGACC 1680  
Db 3681 CACATTTTTCAGGTGTTGCAACACAGGTCTGAGTGTGAGCTCTGGTTCAGCCAGGACC 3740  
Qy 1681 AAAGAGGTGTGAAGTGAAGTGTCTTCAGTNCCTCCAGACATGTGCCCTTTGCTGCTG 1740

Db 3741 AAAGAAGTGTGAAGTGAAGTGTCTCAGT-CCCAGACATGTGCCCTTTTCTGCTG 3799  
Qy 1741 GCTACCACTCTTCCCCAGAGCAGCAGGCCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 1800  
Db 3800 GCTACCACTCTTCCCCAGAGCAGCAGGCCCCGAGCCCTTTCAGGCCAGCACTGCCCCAG 3859  
Qy 1801 ACTGCTGGCACTCAGTTCCTTCATCTCTAAGGTGAAGGTGATGAGGATATGCTCTGA 1860  
Db 3860 ACTGCTGGCACTCAGTTCCTTCATCTCTAAGGTGAAGGTGATGAGGATATGCTCTGA 3919  
Qy 1861 CAGGAACAGTCTGTGATGAGCAGTATCAGTGTCTTAAAGGNAAGCAGCAGAGAGAGCGY 1920  
Db 3920 CAGGAACAGTCTGTGATGAGCAGTATCAGTGTCTTAA--GGAAGCAGCAGAGAGAGCG- 3976  
Qy 1921 TCCGCGCCCCCAGNCCCACACTNATCAGTGTGTCAGGCTGCTGCTGCTGCTGCTGCTG 1980  
Db 3977 TCCGCGCCCCCAGNCCCACACTNATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4032  
Qy 1981 GCTCAGNATCANCACACTGACATCNCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
Db 4033 -----TCAGCATCACACTGACATCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4083  
Qy 2041 CCGNACGGCAGCTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100  
Db 4084 CCG-ACGGCAGCTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4140  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATG 2160  
Db 4141 GGGCAGGGCAGGGCAGTGCACACT-----GTAGGAGCATAGCAGGAGAGATGG 4190  
Qy 2161 GGTGNAAGGNCACAGTCTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2220  
Db 4191 GGTGAA--GGGACACAGTCTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4245  
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCTTTCAGCTGATGATGATGATGATGATGATGATGATGATG 2280  
Db 4246 --CCAGATTTCTTAAGATAGCANCCTTTCAGCTGATGATGATGATGATGATGATGATGATGATG 4301  
Qy 2281 CCGAGGGAGCTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340  
Db 4302 CCGAGGGAGCTTTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4359  
Qy 2341 TGTCTCTAGCTTCAACCTCTTGGGAGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400  
Db 4360 TGTCTCTAGCTTCAACCTCTTGGGAGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4419  
Qy 2401 TCCACATCTTGCACAGACAGCTTTTNGTCCAGCTGTCACATTTAGTGCAGACTGCTCCC 2460  
Db 4420 TCCACATCTTGCACAGACAGCTTTT-GTCCAGCTGTCACATTTAGTGCAGACTGCTCCC 4478  
Qy 2461 GGGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2520  
Db 4479 GGGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4538  
Qy 2521 GGTGTAGAGAACTTTTGTAGAACTTAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 2578  
Db 4539 GGTGTAGAGAACTTTTGTAGAACTTAAAGTTTGGGGTGTAGCAAAATGTTAAAAA 4596

RESULT 12  
US-10-283-975A-294  
; Sequence 294, Application US/10283975A  
; GENERAL INFORMATION:  
; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
; FILE REFERENCE: CDS 293 PCT  
; CURRENT FILING DATE: 2002-10-30  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,938  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/338,997  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,081

; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/341,012  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: HUMAN  
; US-10-283-975A-294

Query Match 70.3%; Score 1850.2; DB 51; Length 4596;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 1 CCGCTCTCACAGCCCGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGGCTGT 60  
Db 2100 CCGCTCTCACAGCCCGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGGCTGT 2159  
Qy 61 CTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGCTGGAAGA 120  
Db 2160 CTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGCTGGAAGA 2219  
Qy 121 GCCCTTGGAGGGAGATATTAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACC 180  
Db 2220 GCCCTTGGAGGGAGATATTAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACC 2279  
Qy 181 ACCAGACCTTCCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGCGCAG 240  
Db 2280 ACCAGACCTTCCATGCCAGCCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGCGCAG 2339  
Qy 241 CTGAGGAGACAAAGGAGGAGCCCTTAAGTTCAGAGCTTCTTCCCAACAGAGCCCCCAG 300  
Db 2340 CTGAGGAGACAAAGGAGGAGCCCTTAAGTTCAGAGCTTCTTCCCAACAGAGCCCCCAG 2399  
Qy 301 AGCCAAACAAGTCTCCTCCCTTGAAGTTCAGAGGAGGAGTCTGGATGTGGAAACCT 360  
Db 2400 AGCCAAACAAGTCTCCTCCCTTGAAGTTCAGAGGAGGAGTCTGGATGTGGAAACCT 2459  
Qy 361 TACCTCTCTCTCTCTGAGCCAGCCCTTGCAGAAACCCAGCTTCCAGAGCGGAAAG 420  
Db 2460 TACCTCTCTCTCTCTGAGCCAGCCCTTGCAGAAACCCAGCTTCCAGAGCGGAAAG 2519  
Qy 421 CAACCGTCCCAGAGAGGAACTGCAGAGCTGGGAATAGATTTTCTCAACAGCCCTGT 480  
Db 2520 CAACCGTCCCAGAGAGGAACTGCAGAGCTGGGAATAGATTTTCTCAACAGCCCTGT 2579  
Qy 481 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC 540  
Db 2580 CCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGCCTCAGCATCGACAGCC 2639  
Qy 541 TCTCCCTGTGGATGACAGTGAAGAAACCCATCAAGGCTTCTCAAGGCTCGCGGACA 600  
Db 2640 TCTCCCTGTGGATGACAGTGAAGAAACCCATCAAGGCTTCTCAAGGCTCGCGGACA 2699  
Qy 601 CCGTACGCTCAGGCGTACACTCTTGGAGCAGCCAGGCTCGAAGCTCCAGCTGGA 660  
Db 2700 CCGTACGCTCAGGCGTACACTCTTGGAGCAGCCAGGCTCGAAGCTCCAGCTGGA 2759  
Qy 661 ACATGCTGTGGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 720  
Db 2760 ACATGCTGTGGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG 2819  
Qy 721 TCCAAATACAGTCTTTAATGTTGAACACCTGTCACATCGGGAGTTCCACCGGGTCAAG 780  
Db 2820 TCCAAATACAGTCTTTAATGTTGAACACCTGTCACATCGGGAGTTCCACCGGGTCAAG 2879  
Qy 781 TGGGAGACATCGCCACTGTCATCAGCAGCCAGATCCAGCTCAGCTTCAAGCTTGGTCA 840  
Db 2880 TGGGAGACATCGCCACTGTCATCAGCAGCCAGATCCAGCTCAGCTTCAAGCTTGGTCA 2939  
Qy 841 CCAAGAGCGGGAGGCTGTTCGCTACGATCGAGTGCAGATCGGGCATCGACCTGC 900





Matches 2334; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

QY	1	CCCCTCTCACAGCCAGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGT	60
Db	2100	CCCCCTCTCACAGCCAGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGT	2159
QY	61	CTGACGCGAGCTGGAGGGAAGGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGA	120
Db	2160	CTGACGCGAGCTGGAGGGAAGGTGAACCCGGGCACTACAGCAAGTGGGAGGTCTGAAGA	2219
QY	121	GCCCTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGGCAATCAAGCCAAATACC	180
Db	2220	GCCCTTGAGGGGAGAAATATAAGAACCAAGACATCCACCGGCAATCAAGCCAAATACC	2279
QY	181	ACCAGACCTCCATGCCAGCCGAGAGCTTTGGCAAGGCGCCACAGGCGCCCGGCCAG	240
Db	2280	ACCAGACCTCCATGCCAGCCGAGAGCTTTGGCAAGGCGCCACAGGCGCCCGGCCAG	2339
QY	241	CTGAGGAGACAAACAGGACAGCCCTTAAGCTCCAGCCTCCTCTCCACACAGAGCCCCAG	300
Db	2340	CTGAGGAGACAAACAGGACAGCCCTTAAGCTCCAGCCTCCTCTCCACACAGAGCCCCAG	2399
QY	301	AGCCAAAACAAGTCTCTCCCTTGAATTTGAGCAAGAGAGTCTGGGATGTGGGAACCT	360
Db	2400	AGCCAAAACAAGTCTCTCCCTTGAATTTGAGCAAGAGAGTCTGGGATGTGGGAACCT	2459
QY	361	TACCTCTGTCTCTGGAGCGAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	420
Db	2460	TACCTCTGTCTCTGGAGCGAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAG	2519
QY	421	CAACCGTCCCGAGCAGGAATGACAGCAGCTGGAATAGAAATTAATCTCAACAGCCTGT	480
Db	2520	CAACCGTCCCGAGCAGGAATGACAGCAGCTGGAATAGAAATTAATCTCAACAGCCTGT	2579
QY	481	CCAGCCATTTTCTCTGAGGAGAGAGGAGCAAAATTTCTCTGTGCTCTAGCATCGACGCC	540
Db	2580	CCAGCCATTTTCTCTGAGGAGAGAGGAGCAAAATTTCTCTGTGCTCTAGCATCGACGCC	2639
QY	541	TCCTCCCTGTCGATGACAGTGAAGAACCCATCAAGAGCTCTCAAGCTCGCGGGACA	600
Db	2640	TCCTCCCTGTCGATGACAGTGAAGAACCCATCAAGAGCTCTCAAGCTCGCGGGACA	2699
QY	601	CCCTGAGCTCAGCGGTACACTCTCTGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGA	660
Db	2700	CCCTGAGCTCAGCGGTACACTCTCTGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGA	2759
QY	661	ACATGGTGTGCGCGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG	720
Db	2760	ACATGGTGTGCGCGGGGGGGCCACCGACACCCCAAGCTATTTCATGGTGTGAAG	2819
QY	721	TCCAAATACAGTCTCTTAATGGTGAACACTTCACATCCGGAGTTCCACCGGGTCAAAG	780
Db	2820	TCCAAATACAGTCTCTTAATGGTGAACACTTCACATCCGGAGTTCCACCGGGTCAAAG	2879
QY	781	TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA	840
Db	2880	TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCA	2939
QY	841	CCAAAGACGGGAGCCTGTTGCTGACGATGGAGGTGCCAGATCGGGCATCGACTGCG	900
Db	2940	CCAAAGACGGGAGCCTGTTGCTGACGATGGAGGTGCCAGATCGGGCATCGACTGCG	2999
QY	901	AGTGACACTGCGCCCTGATGACGCTTCGCTCGAGCTGGAGGTCAAGCATGCGCAGC	960
Db	3000	AGTGACACTGCGCCCTGATGGAGCTTCGCTCGAGCTGGAGGTCAAGCATGCGCAGC	3059
QY	961	TGGAGAACAGGCCCTAACCTGCCCTTCCACCGCGGCTCCACACTGCCGGAAGAGCCT	1020
Db	3060	TGGAGAACAGGCCCTAACCTGCCCTTCCACCGCGGCTCCACACTGCCGGAAGAGCCT	3119
QY	1021	TCTGCTCGGTGACGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGATYTG	1080
Db	3120	TCTGCTCGGTGACGATGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGATYTG	3177
QY	1081	NCCAGCCCCCGGCTCARGNTGGGAACCAAGGCCCTCGMACGACGACGAGTNNAGG 1140	
Db	3178	GCCAGCCCCCGGCTCA - CAGTGGGAACCAAGGCCCTCG - - - - - CAGCAGCAAGTGGGG 3230	
QY	1141	CAAGCAGATGCTCCAGGATTTTCAACNCCTGAGCCNTGCCANCCCTGCTGAADA 1200	
Db	3231	GCAAGCAGATGCTCCAGGATTTTCACTGAGCCCTGCCAC - - - - - CTTGCTGA 3284	
QY	1201	AAACAYTNCGCCACGCTGAGAGACAGAGGATGNCAGGAGTNNACTTGGGAA 1260	
Db	3285	AAAAACATCCGCCACGCTGAGAGACAGAGGATGTC - - - - - AGGATTTACTTGGGAA 3340	
QY	1261	ACAAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTNCAGTTCGCTGNAACCCCTTGG 1320	
Db	3341	ACAAAAACAGGATCTTT - TTCTGCCCCCTGCTCCAGT - CGAGTTGGCCTGA - - CCGCTTG 3396	
QY	1321	ANTCAGTGACATTTTGGCAGANCAGGGGAGAGAGCTTCCAGCTTGGGTGAGAAGG 1380	
Db	3397	GATCAGTGACATTTTGGCAGA - CAGGGGAGAGAGCTTCCAGCTTGGGTGAGAAGG 3455	
QY	1381	GTGGCGAGCCCTTCGGCCCTCACCTTNCAGGCTGCTGTGNAGAGTGTCAAGTGTGA 1440	
Db	3456	GTGGCGAGCCCTTCGGCCCTCACCT - CAGGCTGCTGTG - AGAGTGTCAAGTGTGA 3513	
QY	1441	AGGNCCTCAANCTCAGNTTTCAGTGCAAGAACAGGTCNACAGGATATGCCGCCCGNTA 1500	
Db	3514	AGGCCCAAACTC - - - AGTTTCAGTGCAAGAACAGG - CAGCAGTATGCCGCCCGTAG 3569	
QY	1501	GTTTAAANNNGGGGCTCTNAAACCCCTTGGCTNGGCTNCACCTNGGCGAGCTCANCC 1560	
Db	3570	GT7AA - - - - GGGGGGCTCTAAACCCCTTGGCTNGGCTCA - - - - CTTGGCCAGCTCA - CCC 3621	
QY	1561	CTTTTGGGTGAGGGGAAAGAAATGCTGAGCCCTGGGAAGGCTWCCCTGGTAGAATACAC 1620	
Db	3622	CTTTTGGGTGAGGGGAAAGAAATGCTGAGCCCTGGGAAGGCT - CCTTGGTAGAATACAC 3680	
QY	1621	CACACTTTTTCAGGTTGTTGCAACACAGGCTCTGAGTTGACCTCTGTTTTCAGCAAGGACC 1680	
Db	3681	CACACTTTTTCAGGTTGTTGCAACACAGGCTCTGAGTTGACCTCTGTTTTCAGCAAGGACC 3740	
QY	1681	AAAGAGGTGTAAAGTAAAGTGTTCAGTNCACAGATGTCGCCCTTTCGCTGTG 1740	
Db	3741	AAAGAGGTGTAAAGTAAAGTGTTCAGT - CCCCAGACATGTGCCCCCTTTCGCTGTG 3799	
QY	1741	GCTACACTTTCCTCCAGCAGCAGGCGCCCGAGCCCTTCAGGCCAGCACTGCCCCAG 1800	
Db	3800	GCTACACTTTCCTCCAGCAGCAGGCGCCCGAGCCCTTCAGGCCAGCACTGCCCCAG 3859	
QY	1801	ACTCGCTGGCACTCAGTTTCCCTCATCTGTAAAGGTGAAGGGTGTATGCAAGGATATGCCCTGA 1860	
Db	3860	ACTCGCTGGCACTCAGTTTCCCTCATCTGTAAAGGTGAAGGGTGTATGCAAGGATATGCCCTGA 3919	
QY	1861	CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAGGNAAGACAGCAGAGAGAGCY 1920	
Db	3920	CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAGGNAAGACAGCAGAGAGAGCY - 3976	
QY	1921	TCCGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCTGCTNGGTTNCCCAAGNAGACA 1980	
Db	3977	TCCGCGCCCGCAGNCCCACTNATCAGTGT - - - CAGCTGTCTGTTTCCCAAGNAGACAG - 4032	
QY	1981	GCTNAGNACATCANCACTGACACTNCACTTNGCCCTGCCCTNGGCGCANGAGGTACTG 2040	
Db	4033	-----TCAGCATCACACTGACACTCACCTGCCCTGCCCTGG - - CCAGAGGGTACTG 4083	
QY	2041	CCGNAAGCACTTTCACATCTGATGNAACCTCAAGCACTTTCATGGCTNGCCCTTNG 2100	
Db	4084	CCG - ACAGCACTTTCAGT - - - CTGATGACCTCAAGCACTTTCATGGCTGCCCTCTGGCA 4140	
QY	2101	GCAGGNCAGGNCAGGNCAGTGCACACTGTAGNAGCATANGCAANGCCAGGAGATGG 2160	
Db	4141	GGCAGGCGAGGCGAGTGCACACT - - - - - GTAGGACATACGAAGCCAGGAGATGG 4190	

Qy	2161	GGTGNAAAGGGANCAAGTCTTTGAGCTGTTCANCAATGATGTGACTNCTCAAAACCTCTTN	2222
Db	4191	GGTGAA--GGGACACAGTCTTTGAGCTGTCCA-CATGCAATGTGACT-CCTCAAAACCTCTT-	4245
Qy	2221	NCCAGNATTTCTCTTAAGAAATAGCANCCCCCTTNCCTTCCCATTTGCCCGACGCTTAGCCTCTTCT	2280
Db	4246	--CCAGATTTCTCTAAGAAATAGCACCCCC--TTCCCCATTTGCCCGACCTTAGCCTCTTCT	4301
Qy	2281	CCCAGGGGAGCTANCTCAGGACTCACGTAGCATTTAAATCAGCTGTGNAATCGTCAGGGGG	2340
Db	4302	CCCAGGGGAGCTA-CTCAGGACTCACTGAGCATTTAAATCAGCTGTG-AATCGTCAGGGGG	4359
Qy	2341	TGCTCTGAGCTCAAACTCTCTGGGAGAGGGGACCGCAGACTCCGTGGGAGAAAGCTCAT	2400
Db	4360	TGCTCTGAGCTCAAACTCTCTGGGAGAGGGGACCGCAGACTCCGTGGGAGAAAGCTCAT	4419
Qy	2401	TCCCACATCTTGCCCAAGACAGCCTTTTNGTCCAGCTGTCCACATTTGAGTFCAGACTGTCTCCC	2460
Db	4420	TCCCACATCTTGCCCAAGACAGCCTTTT-GTCCAGCTGTCCACATTTGAGTFCAGACTGTCTCCC	4478
Qy	2461	GGGAGAGAGCCCCGGCCCCCAGCACATAAGAACTCAGCCTTTGGTACTGCGAGAGTCTG	2520
Db	4479	GGGAGAGAGCCCCGGCCCCCAGCACATAAGAACTCAGCCTTTGGTACTGCGAGAGTCTG	4538
Qy	2521	GGTGTGTAGAGAACTCTTTTGTAAAGCAATAAGTTTGGGGTGATGACAAATGTTAAAAAA	2578
Db	4539	GGTGTGTAGAGAACTCTTTTGTAAAGCAATAAGTTTGGGGTGATGACAAATGTTAAAAAA	4596
RESULT 14			
US-60-474-850-144			
; Sequence 144, Application US/60474850			
; GENERAL INFORMATION:			
; APPLICANT: Wohlgemuth, Jay			
; APPLICANT: Lv, Ngoc			
; APPLICANT: Prentice, James			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING			
; TITLE OF INVENTION: RENAL TRANSPLANT REJECTION			
; FILE REFERENCE: 506613000700			
; CURRENT APPLICATION NUMBER: US/60/474,850			
; CURRENT FILING DATE: 2003-05-30			
; NUMBER OF SEQ ID NOS: 552			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 144			
; LENGTH: 4596			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-60-474-850-144			

Query Match	70.3%	Score 1850.2	DB 117	Length 4596
Best Local Similarity	90.5%	Pred. No. 0		
Matches 2334	Conservative 4	Mismatches 159	Indels 81	Gaps 34
Qy 1	CCCTCTCAGAGCCCGAGCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	60		
Db 2100	CCCTCTCAGAGCCCGAGCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT	2159		
Qy 61	CTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTCGGAGGTCCTGAAGA	120		
Db 2160	CTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTCGGAGGTCCTGAAGA	2219		
Qy 121	GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACGCGCCAAATCAAGGCCAATTACC	180		
Db 2220	GCCCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACGCGCCAAATCAAGCCAATTACC	2279		
Qy 181	ACAGACCCCTCCATGCCAGCGGAGAGAGCTTTTCGCCAAGGGCCCCCAGGGCCCCGGCCAG	240		
Db 2280	ACCAGACCCCTCCATGCCCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCCAGGGCCCCGGCCAG	2339		
Qy 241	CTGAGGAGACAACAGGCAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCCCAG	300		
Db 2340	CTGAGGAGACAACAGGCAGAGCCCTTAAGCTCCAGCCTCTCTCTCCACACAGAGCCCCCAG	2399		

QY	301	AGCAAA	CAAGTCTCCTCCTT	GCATTTGAGCAAGAGAGTCTGGATGTGGAAACCCCT	360
DB	2400	AGCAAA	CAAGTCTCCTCCTT	GACTTTGAGCAAGAGAGTCTGGATGTGGAAACCCCT	2459
QY	361	TACCTCTG	TCCCTCGGAGCAGCCCT	TGCCAGAAACCCAGCTCACCAGACGCGAAG	420
DB	2460	TACCTCTG	TCCCTCGGAGCAGCCCT	TGCCAGAAACCCAGCTCACCAGACGCGAAG	2519
QY	421	CAACCGT	CCCGAGCAGGAATCT	CAGCAGCTGGAAATAGAAATTTATTTCTCAACAGAGCTGT	480
DB	2520	CAACCGT	CCCGAGCAGGAATCT	CAGCAGCTGGAAATAGAAATTTATTTCTCAACAGAGCTGT	2579
QY	481	CCGAGCCATTTT	CTCTGGAGGACGAGGAAATTTCTCTGTGCTCTCAGCATCGACAGCC	540	
DB	2580	CCGAGCCATTTT	CTCTGGAGGACGAGGAAATTTCTCTGTGCTCTCAGCATCGACAGCC	2639	
QY	541	TCTCCCTGT	TCGGATGACAGTGAGAAGAACCCCATCAAGGCTCTCAAGCTCGGGACA	600	
DB	2640	TCTCCCTGT	TCGGATGACAGTGAGAAGAACCCCATCAAGGCTCTCAAGCTCGGGACA	2699	
QY	601	CCCTGAGCT	CAGCGGTACACTCTCTGAGCAGCGCAGGCTCGAAGCTCCAGCTTGA	660	
DB	2700	CCCTGAGCT	CAGCGGTACACTCTCTGAGCAGCGCAGGCTCGAAGCTCCAGCTTGA	2759	
QY	661	ACATGTGT	TGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAG	720	
DB	2760	ACATGTGT	TGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAG	2819	
QY	721	TCCAAATACAGTCTCTTAAT	TGTTGAACACTTGACATCCGGAGTTCCACCGGTCNAAG	780	
DB	2820	TCCAAATACAGTCTCTTAAT	TGTTGAACACTTGACATCCGGAGTTCCACCGGTCNAAG	2879	
QY	781	TGGAGACATCGCATCTGGCAT	CAGCAGCAGATCCAGCTGCAGCTTCAGCTTCGTGCTCA	840	
DB	2880	TGGAGACATCGCATCTGGCAT	CAGCAGCAGATCCAGCTGCAGCTTCAGCTTCGTGCTCA	2939	
QY	841	CCAAAGACGGCAGCCTT	TGCTACGACATGAGGTGCCAGACTCGGGCATCGACCTGC	900	
DB	2940	CCAAAGACGGCAGCCTT	TGCTACGACATGAGGTGCCAGACTCGGGCATCGACCTGC	2999	
QY	901	AGTGCA	CACATGGGCCCTTGATGGCAGCTTGCTTGGAGCTTGGAGGTCAAGCATGGCCAGC	960	
DB	3000	AGTGCA	CACATGGGCCCTTGATGGCAGCTTGCTTGGAGCTTGGAGGTCAAGCATGGCCAGC	3059	
QY	961	TGGAGAACAGCCCTTAAC	CCTGCTCCACCGCGCTCCACACTCCCGAAAGCAGCCT	1020	
DB	3060	TGGAGAACAGCCCTTAAC	CCTGCTCCACCGCGCTCCACACTCCCGAAAGCAGCCT	3119	
QY	1021	TCCTGTCT	CGGTGACGATGTGCTTCAAAACACAGGCTCAGCGTTTCCAGGGGATYTG	1080	
DB	3120	TCCTGTCT	CGGTGACGATGTGCTTCAAAACACAGGCTCAGCGTTTCCAGGGGATYTG	3177	
QY	1081	NCAGCCCCCGGCTC	ARCAGNTGGAAACAGGGCTTCGNACGNAGAGTNGGGGG	1140	
DB	3178	GCCAGCCCCCGGCTCA	--CAGTGGGAAACAGGGCTCG-----CAGCAGCAAGTGGGG	3230	
QY	1141	CACGNAGATGCTCC	CAGGATTTCAACNCTTGACCCNCTGCCANCCCTCTCAADA	1200	
DB	3231	GCAAGCAGATGCTCC	CAGGATTTCAACNCTTGACCCNCTGCCANCCCTCTCAADA	3284	
QY	1201	AAACATY	NCGCCACGTGAAGACAGAGGAGTGGNACGGAGTTTNNACCTYGGGGAA	1260	
DB	3285	AAACATY	NCGCCACGTGAAGACAGAGGAGTGGNACGGAGTGGNACCTYGGGGAA	3340	
QY	1261	ACAAAACAGGGATCTTT	NTTCTGCCCTGTTCAGTNCAGATTTGGCTCTGNACCCGCTTGG	1320	
DB	3341	ACAAAACAGGGATCTTT	NTTCTGCCCTGTTCAGTNCAGATTTGGCTCTGNACCCGCTTGG	3396	
QY	1321	ANTCAGTGACCATTTCT	TGTCAGANCAAGGAGCAGCTTCCAGCTCGGTGAGAGGG	1380	
DB	3397	GATCAGTGACCATTTCT	TGTCAGANCAAGGAGCAGCTTCCAGCTCGGTGAGAGGG	3455	
QY	1381	GTGGGCGAGCCCTT	CGGCCCTCACCTTNCAGGCTGTGTGNAGAGTGTCAAGTGTGA	1440	

Db	3456	GTGGGAGCCCTTCGGCCCTCAACCTT-CCAGCTGCTGTG-AGAGTGTCAAGTGTGTA	3513	Db	4479	GGGGAGAGAGCCCGCCCGCCAGACATATAAGAACTGCAGCGCTTGGTACTGCAGAGTCTG	4538
Qy	1441	AGGNCVCAANCTCAGNNTTCAGTGACAGACCAAGTNCAGCAGGTATGCCGCCCGGNTA	1500	Qy	2521	GGTTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAAAA	2578
Db	3514	AGGGCCCAAACTC---AGGTTTCAGTGCAAGAACCAAGGT-CAGCAGGTATGCCGCCCGGTAG	3569	Db	4539	GGTTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAAAA	4596
Qy	1501	GGTTAANNNGGGGCGCTCTNAACCCCTTGCTNGGCTNCACCTNGGCCAGCTTCANCCC	1560	RESULT 15			
Db	3570	GTATA---GGGGGCGCTTAACCCCTTGCTGCGCTCA---CCTGCCAGCTCA-CCC	3621	US-09-496-914A-6881			
Qy	1561	CTTTTGGGTAGGGGAAAAGAAATCCCTGACCCCTGGGAAGCTCCCTGGTGTAGAAATACAC	1620	; Sequence 6881, Application US/09496914A			
Db	3622	CTTTTGGGTAGGGGAAAAGAAATCCCTGACCCCTGGGAAGCT-CCCTGGTGTAGAAATACAC	3680	; GENERAL INFORMATION:			
Qy	1621	CACACTTTTCAGGTGTTGCAACACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC	1680	; APPLICANT: Tang, Yuanhua T.			
Db	3681	CACACTTTTCAGGTGTTGCAACACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACC	3740	; APPLICANT: Tillinghast, John			
Qy	1681	AAAGAAGGTGTGTAAAGTGAAGTGTCTCAGTNCACAGATGTCGCCCTTTGCTGCTG	1740	; APPLICANT: Sinku, Ankura			
Db	3741	AAAGAAGGTGTGTAAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTGCTGCTG	3799	; APPLICANT: Liu, Chenghua			
Qy	1741	GCTACACTCTTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCAGCACTGCCCCAG	1800	; APPLICANT: Drmanac, Radoje T.			
Db	3800	GCTTACACTCTTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCAGCACTGCCCCAG	3859	; TITLE OF INVENTION: Novel Contigs Obtained			
Qy	1801	ACTCGCTGCACTCAGTTCCTCCCTCATCTGTAAAGGTGAAGGTGTATGCAAGTATGCTTGA	1860	; TITLE OF INVENTION: From Various Libraries			
Db	3860	ACTCGCTGCACTCAGTTCCTCCCTCATCTGTAAAGGTGAAGGTGTATGCAAGTATGCTTGA	3919	; FILE REFERENCE: 787			
Qy	1861	CAGGAACAGTCTGTGTAGTGAATGATCATGTCTTAAGNNAAGCAGCAGAGAGAGCY	1920	; CURRENT APPLICATION NUMBER: US/09/496,914A			
Db	3920	CAGGAACAGTCTGTGTAGTGAATGATCATGTCTTAAGNNAAGCAGCAGAGAGAGCG-	3976	; CURRENT FILING DATE: 2000-02-03			
Qy	1921	TCCGGGCCCCAGNCCCCACTNATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1980	; PRIOR APPLICATION NUMBER: US 09/353,690			
Db	3977	TCCGGGCCCCAGNCCCCACTNATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4032	; PRIOR FILING DATE: 1999-07-14			
Qy	1981	GCTNCAGNATCANCACTGACACTNCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2040	; PRIOR APPLICATION NUMBER: US 09/034,341			
Db	4033	-----TCAGATCATACACTGACACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4083	; PRIOR FILING DATE: 1998-02-13			
Qy	2041	CCNAGCGGCACTTGACACTGCTGATGACCTCAAGCACTTTCATGCTGCTGCTGCTGCTGCTG	2100	; PRIOR APPLICATION NUMBER: US 09/045,400			
Db	4084	CCG-ACGGCACTTGGCACT--CTGATGACCTCAAGCACTTTCATGCTGCTGCTGCTGCTGCTG	4140	; PRIOR FILING DATE: 1998-03-20			
Qy	2101	GCAGGNCAGGNCAGGNCAGTGCACANCTGTAGNAGCATANGCAANGCCAGGATGG	2160	; PRIOR APPLICATION NUMBER: US 09/321,214			
Db	4141	GGCAGGGCAGGGCAGTGCACACT-----GTAGGAGCATAGCAAGCCAGGATGG	4190	; PRIOR FILING DATE: 1999-05-26			
Qy	2161	GGTGNAAAGGANCACAGTCTTGAGCTGTCCANCATGATGTGACTTNCCTCAACCTCTTN	2220	; PRIOR APPLICATION NUMBER: US 09/131,598			
Db	4191	GGTGAA--GGGACACAGTCTTGAGCTGTCCA-CATGATGTGACT-CCTCAAACTCTT-	4245	; PRIOR FILING DATE: 1998-08-10			
Qy	2221	NCCAGNATTTCTTAAGAAATAGCANCCCCCTTNCCTTCCCTTCCCTTCCCTTCCCTTCCCTT	2280	; PRIOR APPLICATION NUMBER: US 09/11-01			
Db	4246	--CCAGATTTCTTAAGAAATAGCANCCCC--TTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT	4301	; PRIOR FILING DATE: 1999-06-04			
Qy	2281	CCCAGGGAGGACTNCTCAGGACTCACTGATGATTAATCATGCTGTGNAATCGTCAGGGG	2340	; PRIOR APPLICATION NUMBER: US 09/328,351			
Db	4302	CCCAGGGAGGACTNCTCAGGACTCACTGATGATTAATCATGCTGTGNAATCGTCAGGGG	4359	; PRIOR FILING DATE: 1999-06-04			
Qy	2341	TGCTGTAGCTCAACTCTGCGGAGGGGAGCGGAGACTCCGTTGGGAGAGGCTCAT	2400	; PRIOR APPLICATION NUMBER: US 09/332,782			
Db	4360	TGCTGTAGCTCAACTCTGCGGAGGGGAGCGGAGACTCCGTTGGGAGAGGCTCAT	4419	; PRIOR FILING DATE: 1999-06-14			
Qy	2401	TCCACATCTTGGCAAGACAGCTTTTNGTCCAGCTGTCCATTTGAGTGCAGCTGCTCC	2460	; PRIOR APPLICATION NUMBER: US 09/235,076			
Db	4420	TCCACATCTTGGCAAGACAGCTTTT-GTCCAGCTGTCCATTTGAGTGCAGCTGCTCC	4478	; PRIOR FILING DATE: 1999-01-20			
Qy	2461	GGGGAGAGGCGCGGCCCCCAGACATATAAGAACTGAGCTTGTACTGCAGAGTCTG	2520	; PRIOR APPLICATION NUMBER: US 09/234,611			

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (122)...(2969)  
; OTHER INFORMATION: similar to gil841434 in the genepept database release 114,  
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters  
US-09-496-914A-6881

Query Match 68.1%; Score 1790.4; DB 22; Length 4487;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 2297; Conservative 5; Mismatches 190; Indels 87; Gaps 34;

Qy	1	CCCCTCTCACAGCCAGCCAGCCATCCAAAGAGGGCTTGAGGAAAGAGCCCATCACCGCTGT	60
Db	1993	CCCCCTCTCACAGCCAGCCAGCCATCCAAAGAGGGCTTGAGGAAAGAGCCCATCACCGCTGT	2052
Qy	61	CTGCAGCGGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	120
Db	2053	CTGCAGCGGAGCTGGGAGGAGGTGAACCGGCACTACAGCAAGTGGAGGTCTGAAGA	2112
Qy	121	GCCCTTGAGGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATTACC	180
Db	2113	GCCCTTGAGGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATTACC	2172
Qy	181	ACCAAGCCCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCCAG	240
Db	2173	ACCAAGCCCTCCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGAGGGCCCGGCCAG	2232
Qy	241	CTGAGGAGACAAAGGAGAGAGCCCTTAAGTCTCAGCCTCTCTCCACACGAGCCCGCCAG	300
Db	2233	CTGAGGAGACAAAGGAGAGAGCCCTTAAGTCTCAGCCTCTCTCCACACGAGCCCGCCAG	2292
Qy	301	AGCCAAACAAGTCTCTCCCTTGACTTTTGACAGAGGAGGAGTCTGGAGATGTGGAAACCT	360
Db	2293	AGCCAAACAAGTCTCTCCCTTGACTTTTGACAGAGGAGGAGTCTGGAGATGTGGAAACCT	2352
Qy	361	TACCTCTGTCTCTCTGGAGCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	420
Db	2353	TACCTCTGTCTCTCTGGAGCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	2412
Qy	421	CAACCGTCCGGAGCAGGAATGAGAGCTGGAATATAGAAATTTCTCTCAACAGCCTGT	480
Db	2413	CAACCGTCCGGAGCAGGAATGAGAGCTGGAATATAGAAATTTCTCTCAACAGCCTGT	2472
Qy	481	CCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC	540
Db	2473	CCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC	2532
Qy	541	TCTCCCTGTCTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGACA	600
Db	2533	TCTCCCTGTCTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGCTCGCGGACA	2592
Qy	601	CCCTGAGCTCAGGCGTACATCTCTGGAGAGCAGGCGGAGGCTCGAAGTCTCAGCTGGA	660
Db	2593	CCCTGAGCTCAGGCGTACATCTCTGGAGAGCAGGCGGAGGCTCGAAGTCTCAGCTGGA	2652
Qy	661	ACATGGTCTGGCCGGGGGGGGCCCGACACACCCCAAGCTATTTCATGGTGTGAAG	720
Db	2653	ACATGGTCTGGCCGGGGGGGGCCCGACACACCCCAAGCTATTTCATGGTGTGAAG	2712
Qy	721	TCCAAATACAGTCTTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGTCAAAG	780
Db	2713	TCCAAATACAGTCTTTAATGGTGAACACCTGACATCCCGGAGTTCCACCGGTCAAAG	2772
Qy	781	TGGAGACATCGGCATCGGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGCTTGCTCA	840
Db	2773	TGGAGACATCGGCATCGGCATCAGCAGCCAGATCCAGCTGCAAGCTTCAGCTTGCTCA	2832
Qy	841	CCAAAGAGCGGCACCTTTGCTTACCATGGAGGTGCCAGACTCGGGCATCGACCTGC	900
Db	2833	CCAAAGAGCGGCACCTTTGCTTACCATGGAGGTGCCAGACTCGGGCATCGACCTGC	2892

Qy	901	AGTGCAACATCGGCCCTTGATGGCAGCTTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGC	960
Db	2893	AGTGCAACATCGGCCCTTGATGGCAGCTTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGC	2952
Qy	961	TGAGAAACAGCGCCCTAACCTGCGCTCCACCGCGCTCCACACTGCGGAAAGACAGCCT	1020
Db	2953	TGAGAAACAGCGCCCTAACCTGCGCTCCACCGCGCTCCACACTGCGG--AAGCAGCCT	3011
Qy	1021	TCCTGCTCGGTGACCATGTGCTGAAACACAGGCTCAGCCGTTTCCAGGGGATVGT	1080
Db	3012	TCCTGCTCGGTGACCATGTGCTGAAACACAGGCTCAGCCGTTTCCAGGG--ATC	3067
Qy	1081	NCAGAGCCCGCGGTCTACAGNTGGGAAACAGAGGCTCTGCNACAGNAGNAGTNGGGG	1140
Db	3068	TGCAGAGCCCGCGGTCTCAGAGTGGGACAGGGCTCTGCAGCAGCAAG-----GTGGG	3120
Qy	1141	CAAGCAGAGTGTCTCCAGGATTTACANCTTGAGCCNTGCCANCCCTGCTGAGADA	1200
Db	3121	GCAAGCAGAGTGTCTCCAGGATTTTACACC---TGAGCCCTGCCCCACCCCTGCTGAGAA	3177
Qy	1201	AAACAYTNCCGCCACGTGAAGAGACAGAGGAGTGCNACAGGAGTTNNACCTTGGGGAA	1260
Db	3178	AAACT---CCGCCACGTGAAGAGACA---GAGAGGATGGCAGGAGTTACCTCGGGAAA	3231
Qy	1261	ACAAAACAGGGATCTTTNTTCTGCCCCCTGCTCCAGTNCAGTGGCTGAGACCCGCTGG	1320
Db	3232	CAAAACAGGATCTTCTGCTGCTGCTCCAG-----TCGAGTTGGCTGAGCCGCTT--G	3282
Qy	1321	ANTCAGTGACCATTTGTTGGCAGANCAGGGGAGAGAGCTTCCAGCTTGGGTGAGAAGGG	1380
Db	3283	GATCAGTGACCATTTGTTGGCAGA-CAGGGGAGAGAGCTTCCAGCTTGGGTGAGAAGGG	3341
Qy	1381	GTGGGAGAGCCTTCGGCCCTCAACCTTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA	1440
Db	3342	GTGGGAGAGCCTTCGGCCCTCAACCT--CAAGCTGCTGTG--AGAGTGTCAAGTGTGTA	3399
Qy	1441	AGGNCACAAANTCAGGNTTCAGTGCAGAACACAGGTCNACAGAGGTATGCCGCCCGNTA	1500
Db	3400	AGGGCCCAAACTC---AGGTTCAGTGCAGAACACAGT-CAGCAGGTATGCCGCCCGTAG	3455
Qy	1501	GGTTAANNNGGGGCCCTCTNAAACCCCTTGCTTNGGCTTNCACCTNGGCCAGCTCANCCC	1560
Db	3456	GTATA---GGGGGCCCTCTAATAACCCCTTGCTTGGCCCTCA---CCTGGCCAGCTCA-CCC	3507
Qy	1561	CTTTTGGGTGTAGGGGAAAGAAATGCTGACCTCGGGAAGGCTTWCCTCTGGTAGAATAAC	1620
Db	3508	CTTTTGGGTGTAGGGGAAAGAAATGCTGACCTCGGGAAGGCT-CCCTGGTAGAATAAC	3566
Qy	1621	CACACTTTTTCAGGTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTCAGCCAGGACC	1680
Db	3567	CACACTTTTTCAGGTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTCAGCCAGGACC	3626
Qy	1681	AAAGAGGTGTGTAAGTGAAGTGTCTGATNCCCAGACATGTGCCCTTCTGCTGCTG	1740
Db	3627	AAAGAGGTGTGTAAGTGAAGTGTCTGAGT-CCCCAGACATGTGCCCTTCTGCTGCTG	3685
Qy	1741	GCTTACCACCTTCCCCAGAGCAGCGCCCGAGCCCTTTCAGSCCCAGACACTGCCCCAG	1800
Db	3686	GCTTACCACCTTCCCCAGAGCAGCGCCCGAGCCCTTTCAGSCCCAGACACTGCCCCAG	3745
Qy	1801	ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGATGAGGATATGCTGTA	1860
Db	3746	ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGATGAGGATATGCTGTA	3805
Qy	1861	CAGGAAACAGTCTGTGATGAGCATGATCAGTGTNAAAGNAAAGCAGCAGAGAGAGGY	1920
Db	3806	CAGGAAACAGTCTGTGATGAGCATGATCAGTGTGTA--GGAAAGCAGCAGAGAGAGCGC	3863
Qy	1921	TCCGGCGCCCGCCAGNCCCCTNATCAGTGTNCCAGGCTGTCTGCTTNCCTCCAGNAGACA	1980
Db	3864	TCCGGCGCCCGCCAGNCCCCTNATCAGTGT---CCAGGCTGTGCTTCC---CAGA	3912
Qy	1981	GCTNCAAGNATCANCACTNCACTNCCCTGCTGCCCTNGGCCANGAGGCTACTG	2040



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 1006.45 Seconds  
(without alignments)  
11471.453 Million cell updates/sec

Title: US-09-155-676B-3  
Perfect score: 2631  
Sequence: 1 cccctctcacagccagcc.....accgtgacctcgaggsgggg 2631

Scoring table: IDENTITY NUC  
Gapop 10'0 , Gapext 1.0

Searched: 12985247 seqs, 2194111492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New:\*

1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:.\*  
2: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:.\*  
3: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:.\*  
4: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:.\*  
5: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:.\*  
6: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:.\*  
7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:.\*  
8: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:.\*  
9: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:.\*  
10: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:.\*  
11: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:.\*  
12: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq:.\*  
13: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq:.\*  
14: /cgn2\_6/ptodata/2/pna/US11\_NEW\_COMB.seq:.\*  
15: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Length	DB ID	Description
1	2553.8	97.1	2631	6 US-09-155-676B-3	Sequence 3, Appli
2	1851.2	70.4	4596	6 US-09-155-676B-6	Sequence 6, Appli
3	1850.2	70.3	4596	8 US-10-450-763-15294	Sequence 15294, A
4	1781.4	67.7	4458	15 US-60-659-397-567	Sequence 567, App
5	1781.4	67.7	4684	15 US-60-659-397-566	Sequence 566, App
6	971.8	36.9	65967	15 US-60-659-397-12129	Sequence 12129, A
7	769.2	29.2	1400	13 US-11-060-756-560	Sequence 560, App
8	769.2	29.2	1400	13 US-11-060-756-560	Sequence 4832, App
9	691.8	26.3	2760	8 US-10-450-763-15295	Sequence 15295, A
10	588.2	22.4	722	8 US-10-450-763-15296	Sequence 15296, A
11	295.6	11.2	3658	15 US-60-680-544-32424	Sequence 32424, A
12	295.6	11.2	3658	15 US-60-680-473-32424	Sequence 32424, A
13	257.4	9.8	761	15 US-60-680-544-10785	Sequence 10785, A
14	257.4	9.8	761	15 US-60-680-544-11056	Sequence 11056, A
15	257.4	9.8	761	15 US-60-680-473-10785	Sequence 10785, A
16	257.4	9.8	761	15 US-60-680-473-11056	Sequence 11056, A
17	200.6	7.6	201	15 US-60-659-397-7882	Sequence 7882, App
18	200.6	7.6	201	15 US-60-659-397-7887	Sequence 7887, App

19 200.6 7.6 201 15 US-60-659-397-7888 Sequence 7888, Ap  
20 200.6 7.6 201 15 US-60-659-397-7893 Sequence 7893, Ap  
21 189.6 7.2 201 15 US-60-659-397-7885 Sequence 7885, Ap  
22 189.6 7.2 201 15 US-60-659-397-7891 Sequence 7891, Ap  
23 189.6 7.2 201 15 US-60-659-397-32724 Sequence 32724, A  
24 188.6 6.4 201 15 US-60-659-397-7886 Sequence 7886, Ap  
25 188.6 6.4 201 15 US-60-659-397-7892 Sequence 7892, Ap  
26 168.6 6.4 201 15 US-60-659-397-32758 Sequence 32758, A  
27 150 5.7 201 15 US-60-659-397-32751 Sequence 32751, A  
28 128.6 4.9 201 15 US-60-659-397-32755 Sequence 32755, A  
29 119.2 4.5 1353 15 US-60-680-544-37753 Sequence 37753, A  
30 119.2 4.5 1353 15 US-60-680-473-37753 Sequence 37753, A  
31 96.4 3.7 201 15 US-60-659-397-7884 Sequence 7884, Ap  
32 96.4 3.7 201 15 US-60-659-397-7890 Sequence 7890, Ap  
33 96.4 3.7 201 15 US-60-659-397-7857 Sequence 7857, A  
34 88 3.3 201 15 US-60-659-397-32726 Sequence 32726, A  
35 78 3.0 201 15 US-60-659-397-32752 Sequence 32752, A  
36 78 3.0 599 9 US-10-972-079-7876 Sequence 7876, Ap  
37 70 2.7 1024 8 US-10-450-763-15297 Sequence 15297, A  
38 68.8 2.6 201 15 US-60-659-397-32718 Sequence 32718, A  
39 62.4 2.4 600 9 US-10-972-079-7877 Sequence 7877, Ap  
40 59.4 2.3 493 1 PCT-US05-00517-813 Sequence 813, App  
41 53.2 2.0 1253 6 US-09-155-676B-4 Sequence 4, Appli  
42 53.2 2.0 1253 13 US-11-060-756-1553 Sequence 1553, Ap  
43 53.2 2.0 1253 13 US-11-060-756-1554 Sequence 1554, Ap  
44 53.2 2.0 1253 13 US-11-060-756-5825 Sequence 5825, Ap  
45 53.2 2.0 1253 13 US-11-060-756-5826 Sequence 5826, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-155-676B-3

; Sequence 3, Application US/09155676B

; GENERAL INFORMATION: WALLACH, David

; APPLICANT: MALININ, Nikolai

; APPLICANT: BOLDIN, Mark

; APPLICANT: KOVALENKO, Andrei

; APPLICANT: METT, Igor

; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR

; FILE REFERENCE: WALLACH-21

; CURRENT APPLICATION NUMBER: US/09/155.676B

; PRIOR FILING DATE: 1999-01-04

; PRIOR APPLICATION NUMBER: PCT/IL97/00117

; PRIOR FILING DATE: 1997-04-01

; PRIOR APPLICATION NUMBER: IL 117800

; PRIOR FILING DATE: 1996-04-02

; PRIOR APPLICATION NUMBER: IL 119133

; PRIOR FILING DATE: 1996-08-26

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 3

; LENGTH: 2631

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1081)..(1081)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1102)..(1102)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1120)..(1120)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1125)..(1125)

[illegible]



; NAME/KEY: misc feature									
; LOCATION: (1984)..(1984)									
; OTHER INFORMATION: n is a, c, g, or t									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (1988)..(1988)									
; OTHER INFORMATION: n is a, c, g, or t									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (1994)..(1994)									
; OTHER INFORMATION: n is a, c, g, or t									
; FEATURE:									
; NAME/KEY: misc feature									
; LOCATION: (2005)..(2005)									
; OTHER INFORMATION: n is a, c, g, or t									
Query Match 97.1%; Score 2553.8; DB 6; Length 2631;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CCCCTCTCACAGCCCGGCCATCCAAAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT	60						
Db	1	CCCCTCTCACAGCCCGGCCATCCAAAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT	60						
Qy	61	CTGCAGCGGAGCTGGGAGGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA	120						
Db	61	CTGCAGCGGAGCTGGGAGGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA	120						
Qy	121	GCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC	180						
Db	121	GCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC	180						
Qy	181	ACCAGACCTCTCATGCCCGAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCCAG	240						
Db	181	ACCAGACCTCTCATGCCCGAGCCGAGAGAGCTTTCCGCAAGGGCCCGAGGGCCCGGCCAG	240						
Qy	241	CTGAGGAGACAACAGGCGAGAGCCCTTAAGTCTCAGCTCTCTCCCAACAGAGCCCGCCAG	300						
Db	241	CTGAGGAGACAACAGGCGAGAGCCCTTAAGTCTCAGCTCTCTCCCAACAGAGCCCGCCAG	300						
Qy	301	AGCCAAACAGTCTCTCCCTTGACTTTGACAGAGGAGGAGTCTGGGATGTGGGAACCCCT	360						
Db	301	AGCCAAACAGTCTCTCCCTTGACTTTGACAGAGGAGGAGTCTGGGATGTGGGAACCCCT	360						
Qy	361	TACCTCTGTCTCTCCCTGGAGCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	420						
Db	361	TACCTCTGTCTCTCTGGAGCCAGCCCTCTGCAGAAACCCAGCTCACAGAGCGGAAAG	420						
Qy	421	CAACCGTCCCGAGCAGAACTGCAGCAGCTGGAAATAGAAATATTCTCAACAGCCCTGT	480						
Db	421	CAACCGTCCCGAGCAGAACTGCAGCAGCTGGAAATAGAAATATTCTCAACAGCCCTGT	480						
Qy	481	CCAGACCAATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTGCTCAGCATGCAGAGCC	540						
Db	481	CCAGACCAATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTGCTCAGCATGCAGAGCC	540						
Qy	541	TCTCCCTGTGCGATGACAGTGAGAAAGAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA	600						
Db	541	TCTCCCTGTGCGATGACAGTGAGAAAGAACCATCAAAAGGCTCTCAAAAGCTCGCGGACA	600						
Qy	601	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA	660						
Db	601	CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGA	660						
Qy	661	ACATGGTGTGGCCCGGGGGCCCAACGACACCCCAAGCTATTTCATAGTGGTGAAG	720						
Db	661	ACATGGTGTGGCCCGGGGGCCCAACGACACCCCAAGCTATTTCATAGTGGTGAAG	720						
Qy	721	TCCAAATACAGTCTCTTAATGGTGAACACCTTCCGGAGTTCACCGGCTCAAG	780						
Db	721	TCCAAATACAGTCTCTTAATGGTGAACACCTTCCGGAGTTCACCGGCTCAAG	780						
Qy	781	TGGGAGACATCGGCATCAGCAGCCAGATCCCGAGCTTCAGCTTGGTCA	840						
Db	781	TGGGAGACATCGGCATCAGCAGCCAGATCCCGAGCTTCAGCTTGGTCA	840						
Qy	841	CCAAAGACGGCAGCCTGTTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC	900						
Db	841	CCAAAGACGGCAGCCTGTTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGC	900						
Qy	901	AGTGACACTGGCCCTGTATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC	960						
Db	901	AGTGACACTGGCCCTGTATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGC	960						
Qy	961	TGGAGAACGGCCCTAACCTCGCCCTCCAGCGCGGCTCCACACTGCGGGAAGCAGCCT	1020						
Db	961	TGGAGAACGGCCCTAACCTCGCCCTCCAGCGCGGCTCCACACTGCGGGAAGCAGCCT	1020						
Qy	1021	TCCTGCTCGGTGACAGATGCTGCCCTGAAAACACAGGCTCAGCCGTTCCCAAGGGATYTG	1080						
Db	1021	TCCTGCTCGGTGACAGATGCTGCCCTGAAAACACAGGCTCAGCCGTTCCCAAGGGATYTG	1080						
Qy	1081	NCCAGCCCGCGGCTCARCAGNTGGGAAACAGGGCTCGNCGAGNAGCNAAAGTNGGGG	1140						
Db	1081	NCCAGCCCGCGGCTCARCAGNTGGGAAACAGGGCTCGNCGAGNAGCNAAAGTNGGGG	1140						
Qy	1141	CAAGCNAGATGCTCCAGGATTTACANCTGAGCCGCTCCCGCCANCCCTGCTGAADA	1200						
Db	1141	CAAGCNAGATGCTCCAGGATTTACANCTGAGCCGCTCCCGCCANCCCTGCTGAADA	1200						
Qy	1201	AAACAYTNCGCCACAGTGAAAGAGACAGAGGATGNCAGAGGATNNACCTYGGGGAA	1260						
Db	1201	AAACAYTNCGCCACAGTGAAAGAGACAGAGGATGNCAGAGGATNNACCTYGGGGAA	1260						
Qy	1261	ACAAAACAGGATCTTTNTTCTGCCCCCTGCTCCAGTNCAGTTCGCCCTGNAACCCGCTTG	1320						
Db	1261	ACAAAACAGGATCTTTNTTCTGCCCCCTGCTCCAGTNCAGTTCGCCCTGNAACCCGCTTG	1320						
Qy	1321	ANTCAGTGACATTTGTTGGCAGNACAGGGGAGAGAGCTTCAGCCCTGGGTCAAGAGGG	1380						
Db	1321	ANTCAGTGACATTTGTTGGCAGNACAGGGGAGAGAGCTTCAGCCCTGGGTCAAGAGGG	1380						
Qy	1381	GTGGGCGAGCCCTTCGGCCCTCACCTTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA	1440						
Db	1381	GTGGGCGAGCCCTTCGGCCCTCACCTTNCAGGCTCTGTGNAGAGTGTCAAGTGTGTA	1440						
Qy	1441	AGGNCCTCAANCTCAGGNTTCAGTGAGAAACAGGATNCAGCAGGTATGCCGCCCGNTA	1500						
Db	1441	AGGNCCTCAANCTCAGGNTTCAGTGAGAAACAGGATNCAGCAGGTATGCCGCCCGNTA	1500						
Qy	1501	GGTTAANNNGGGGCTCTNAAACCCCTTGGCTNGGCTNCACTNGGCGCAGCTCANCCC	1560						
Db	1501	GGTTAANNNGGGGCTCTNAAACCCCTTGGCTNGGCTNCACTNGGCGCAGCTCANCCC	1560						
Qy	1561	CTTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCTWCCCTGGTAGAATACAC	1620						
Db	1561	CTTTTGGGTGTAGGGGAAAAGATGCTGACCTTGGGAAGGCTWCCCTGGTAGAATACAC	1620						
Qy	1621	CACACTTTTTCAGGTTGTTCACACACAGGCTCTGAGTTGACCTCTGGTTCAAGCAAGACC	1680						
Db	1621	CACACTTTTTCAGGTTGTTCACACACAGGCTCTGAGTTGACCTCTGGTTCAAGCAAGACC	1680						
Qy	1681	AAAGAGGTGTGAAGTGAAGTGTCTCATGATNCCCGACACATGTGCCCCCTTGTGCTGTG	1740						
Db	1681	AAAGAGGTGTGAAGTGAAGTGTCTCATGATNCCCGACACATGTGCCCCCTTGTGCTGTG	1740						
Qy	1741	GCTTACCACTTCTCCAGAGCAGGCGCCCGAGCCCTTTCAGGCCCGCAGCACTGCCCCAG	1800						
Db	1741	GCTTACCACTTCTCCAGAGCAGGCGCCCGAGCCCTTTCAGGCCCGCAGCACTGCCCCAG	1800						
Qy	1801	ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGAGGATATGCTCTGA	1860						
Db	1801	ACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGTGAAGGTGATGAGGATATGCTCTGA	1860						
Qy	1861	CAGGAACAGTCTGTGGATGACATGATCAGTGTCTNAAAGGNAAGCAGCAGAGAGAGGY	1920						
Db	1861	CAGGAACAGTCTGTGGATGACATGATCAGTGTCTNAAAGGNAAGCAGCAGAGAGAGGY	1920						





QY	61	CTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTCTGAAGA	120	Db	3231	GCACACAGATGCCTCCAGGATTTCAACCTGAGCCCTGCCAC	3284
Db	2160	CTGCAGCGGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTCTGAAGA	2219	QY	1201	AAACATYTCGCCACCGTGAAGAGACAGAGAGTGNACAGGATTNNACTCTGGGAA	1260
QY	121	GCCCTTGGAGGGGAGAAATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC	180	Db	3285	AAAAATCATCGGCACCGTGAAGAGACAGAGAGGAGGATGCG	3340
Db	2220	GCCCTTGGAGGGGAGAAATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATACC	2279	QY	1261	ACAAAACAGGATCTTTTCTGCCCCCTCTCCAGTNCAGTTCGCTGNACCCCTCTGG	1320
QY	181	ACCAGACCTTCATGCCCGCAGCGAGAGAGCTTTCGCAAGGCCCCAGGGCCCGGCCAG	240	Db	3341	ACAAAACAGGAGTCTTTCTGCCCCCTCTCCAGTNCAGTTCGCTGNACCCCTCTGG	3396
Db	2280	ACCAGACCTTCATGCCCGCAGCGAGAGAGCTTTCGCAAGGCCCCAGGGCCCGGCCAG	2339	QY	1321	AMTCACTGACCAATTTGTTGGCAGANACAGGGAGACAGCTTCACGCTGGGTCAAGGG	1380
QY	241	CTGAGACACAACAGGCGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCGCAG	300	Db	3397	GATCAGTGACCAATTTGTTGGCAGA-CAGGGAGAGCAGCTTCAGCCTGGGTCAAGAGG	3455
Db	2340	CTGAGAGACACAAGGCGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAGCCCGCAG	2399	QY	1381	GTGGGCGAGCCCTTGGCCCCCTCACCTTNCAGGCTGTGTGNAGAGTGTCAAGTGTGTA	1440
QY	301	AGCCAAAACAAGTCTCCTCCTTGAATTTGACAGAGGAGGTCTGGGATGTGGGAAACCT	360	Db	3456	GTGGGCGAGCCCTTGGCCCCCTCACCTTNCAGGCTGTGTGNAGAGTGTCAAGTGTGTA	3513
Db	2400	AGCCAAAACAAGTCTCCTCCTTGAATTTGACAGAGGAGGTCTGGGATGTGGGAAACCT	2459	QY	1441	AGGNCCTCAAACTCAGGNTTCAGTGCAGAAACAGGTNCAGCAGGTATGCCCGCCGNTA	1500
QY	361	TACCTCTGTCTCCTCGAGCGAGCCCTGCGAGAAACCCAGCTCCACAGAGCGGAAAG	420	Db	3514	AGGCCCCAAATC--AGGTTCACTGCAGAACAGGT-CAGCAGGTATGCCCGCCGTAG	3569
Db	2460	TACCTCTGTCTCCTCGAGCGAGCCCTGCGAGAAACCCAGCTCCACAGAGCGGAAAG	2519	QY	1501	GTTAANNNGGGCCCTCTNAAACCCCTTGGCTTGCCTTGCCTCA--CCTGGCCAGCTCA	3621
QY	421	CAACCGTCCCGAGCAGGAATCTGACAGAGCTGGGAAATAGAAATTTCTCTCAACAGCCTGT	480	Db	1561	CTTTTGGGTGAGGGGAAAGAAATGCTCACTTGGGAAAGGCTTCCCTTGGTGAATACAC	1620
Db	2520	CAACCGTCCCGAGCAGGAATCTGACAGAGCTGGGAAATAGAAATTTCTCTCAACAGCCTGT	2579	QY	3622	CTTTTGGGTGAGGGGAAAGAAATGCTCACTTGGGAAAGGCT-CCCTGGTGAATACAC	3680
QY	481	CCGAGCAATTTCTCTGGAGGAGCAGGAGCAATTTCTCTGTCCTCAGCATCCAGCAGCC	540	Db	1621	CACACTTTTTCAGGTTGTTGCAACACAGGTCTCAGTTGACCTCTGTTTCAGCCAGGACC	1680
Db	2580	CCGAGCAATTTCTCTGGAGGAGCAGGAGCAATTTCTCTGTCCTCAGCATCCAGCAGCC	2639	QY	3681	CACACTTTTTCAGGTTGTTGCAACACAGGTCTCAGTTGACCTCTGTTTCAGCCAGGACC	3740
QY	541	TCTCCTCTGCGATGACAGTGAGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACA	600	Db	1681	AAAGAGGTGTAAAGTGAAGTGTCTCAGTNCACAGACATGTGCCCTTTGCTGCTG	1740
Db	2640	TCTCCTCTGCGATGACAGTGAGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACA	2699	QY	3741	AAAGAGGTGTAAAGTGAAGTGTCTCAGTNCACAGACATGTGCCCTTTGCTGCTG	3799
QY	601	CCCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTCCAGCTGGA	660	Db	1741	GCTACCACTCTTCCCGAGAGCAGCGCCCGAGCCCTTCAGGCCAGACACTGCCCCAG	1800
Db	2700	CCCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGCGCGAGGCTCGAAGCTCCAGCTGGA	2759	QY	3800	GCTACCACTCTTCCCGAGAGCAGCGCCCGAGCCCTTCAGGCCAGACACTGCCCCAG	3859
QY	661	ACATGTGTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCATAGTGTGTAAG	720	Db	1801	ACTCGCTGGCACTCAGTTCCCTCACTGTAAAGTGAAGGTGATGCAGGATATGCTGA	1860
Db	2760	ACATGTGTGCGCGCGCGCGCCACCGACACCCCAAGCTATTTCATAGTGTGTAAG	2819	QY	3860	ACTCGCTGGCACTCAGTTCCCTCACTGTAAAGTGAAGGTGATGCAGGATATGCTGA	3919
QY	721	TCCAAATACAGTCTCTTAATGTTGAACAACCTGACATCCGGAGTTCACCGGTTCAAG	780	Db	1861	CAGNAACAGTCTGTGGATGGACATGATCAGTGTCTAAAGAGGAGGAGAGAGAGCGY	1920
Db	2820	TCCAAATACAGTCTCTTAATGTTGAACAACCTGACATCCGGAGTTCACCGGTTCAAG	2879	QY	3920	CAGNAACAGTCTGTGGATGGACATGATCAGTGTCTAAAGAGGAGGAGAGAGAGCGY	3976
QY	781	TGGGAGACATCGCACTGGCATCAGACGACAGATCCAGCTGCAGCTTCAGCTTGGTCA	840	Db	1921	TCCGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTG	1980
Db	2880	TGGGAGACATCGCACTGGCATCAGACGACAGATCCAGCTGCAGCTTCAGCTTGGTCA	2939	QY	3977	TCCGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTG	4032
QY	841	CCAAAGACGGCGAGCCTGTTTCGCTACGACATGAGGTGCGAGACTCGGGGATCGACCTGC	900	Db	1981	GCTNAGNATCANCATGACACTGACACTNCCCTGCCCCCTGCCCCCTGCCCCCTGCCCC	2040
Db	2940	CCAAAGACGGCGAGCCTGTTTCGCTACGACATGAGGTGCGAGACTCGGGGATCGACCTGC	2999	QY	4033	-----TCAGCATCAGACTGACACTCACCCTGCCCCCTGCCCCCTGCCCCCTGCCCC	4083
QY	901	AGTGCACTATGCGCCCTGATGCGAGCTTCGCTGAGCTGGAGGTCAGCATGCGCAGC	960	Db	2041	CCGNAACGCGACTTTTGACNTCTGATGNACTCAAAGCACTTTTCATGCTGCTGCTGCTG	2100
Db	3000	AGTGCACTATGCGCCCTGATGCGAGCTTCGCTGAGCTGGAGGTCAGCATGCGCAGC	3059	QY	4084	CCG-ACGGCACTTTGCACT--CTGATGACCTCAAAGCACTTTTCATGCTGCTGCTGCTG	4140
QY	961	TGGAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT	1020	Db	2101	GCAGGNCAGGNCAGGNCAGTGCACACTGTAGGNACATANGCAANGCCAGGAGATGG	2160
Db	3060	TGGAGAACAGGCCCTTAACCTTGCCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCT	3119	QY	4141	GGGCGAGGCGAGGCGAGTGACACT-----GTAGGAGCATAGCAGCCAGGAGATGG	4190
QY	1021	TCTGCTCGGTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080	Db	2161	GGTGAAGGAGNACAGTCTTTCAGCTGTCCANCATGTCATGTCATCTCAAACCTCTTN	2220
Db	3120	TCTGCTCGGTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3177	QY	4191	GGTGAAG-GGGACACAGTCTTTCAGCTGTCCA-CATGATGTGACT-CCTCAAACCTCTT-	4245
QY	1081	NCCAGCCCCCGGCTCARGAGTGGGAACAGAGGCCCTCGNAGCAGNAGAGGTTGGGG	1140	Db	2221	NCCAGNATTTCTTAAGAAATAGCAGNCCCTTNCCTTCCCATGCCCCAGCTTACCTCTTCT	2280
Db	3178	GCCAGCCCCCGGCTCARGAGTGGGAACAGAGGCCCTCGNAGCAGNAGAGGTTGGGG	3230	QY			
QY	1141	CAAGCAGAAATGCTTCCAGGATTTTCANCTGAGCCCTTGCCTGCTGAADA	1200				

Db 4246 --CCAGATTCTCTAAGATAGACCCCC--TTCCCCATTCGCCAGCTTAGCCTCTTCT 4301  
Qy 2281 CCCAGGAGACTANTCAGGACTCAGCTAGCATTAATCAGCTGNAATCGTCAGGGGG 2340  
Db 4302 CCCAGGAGACTA-CTCAGGACTCAGCTAGCATTAATCAGCTGTG-AATCGTCAGGGGG 4359  
Qy 2341 TGTCTGTAGCTCAACTCTCTGGGCGAGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 2400  
Db 4360 TGTCTGTAGCTCAACTCTCTGGGCGAGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 4419  
Qy 2401 TCCACATCTTGGCAAGACAGCTTTTNGTCCAGCTGTCCACATTAAGTTCAGCTGCTCCC 2460  
Db 4420 TCCACATCTTGGCAAGACAGCTTTT-GTCCAGCTGTCCACATTAAGTTCAGCTGCTCCC 4478  
Qy 2461 GGGGAGAGAGCCCGGCCCCCAGACACATAAAGAACTGCAGCTTTGGTGTGATGACAAATGTTAAAAA 2520  
Db 4479 GGGGAGAGAGCCCGGCCCCCAGACACATAAAGAACTGCAGCTTTGGTGTGATGACAAATGTTAAAAA 4538  
Qy 2521 GGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 2578  
Db 4539 GGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGTATGACAAATGTTAAAAA 4596

RESULT 4  
US-60-659-397-567  
; Sequence 567, Application US/60659397  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: CHANG, Sheng-Yung  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001470  
; CURRENT APPLICATION NUMBER: US/60/659,397  
; CURRENT FILING DATE: 2005-03-09  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 567  
; LENGTH: 4458  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-659-397-567

Query Match 67.7%; Score 1781.4; DB 15; Length 4458;  
Best Local Similarity 89.1%; Pred. No. 0;  
Matches 2292; Conservative 10; Mismatches 183; Indels 87; Gaps 35;

Qy 1 CCCCTCTCACAGCCAGGCCATCCAGAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 60  
Db 1974 CCCCTCTCACAGCCAGGCCATCCAGAGAGGGGCTGAGGAAAGAGCCCATCCACCGCGTGT 2033  
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 120  
Db 2034 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGA 2093  
Qy 121 GCCCTTGGAGGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATTACC 180  
Db 2094 GCCCTTGGAGGGGAGATATAAGAAACAAGACATCCACCGCCAAATCAAGCCAAATTACC 2153  
Qy 181 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGGAGGGCCCGGCGCAG 240  
Db 2154 ACCAGACCTTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCGGAGGGCCCGGCGCAG 2213  
Qy 241 CTGAGGAGACAAAGGAGAGCCCTTAAGTCTCAGCTTCCAGCTTCTCCACCGAGAGCCCGCAG 300  
Db 2214 CTGAGGAGACAAAGGAGAGCCCTTAAGTCTCAGCTTCCAGCTTCTCCACCGAGAGCCCGCAG 2273  
Qy 301 AGCCAAACAAGTCTCTCCCTTGAATTTGACAGAGGAGGAGTCTGGGATGTGGAAACCTT 360  
Db 2274 AGCCAAACAAGTCTCTCCCTTGAATTTGACAGAGGAGGAGTCTGGGATGTGGGAAACCTT 2333  
Qy 361 TACCTCTGCTCTCCCTGGAGGCCGCCCTTGCAGAAACCCAGCTCACCAGAGCGGAAG 420

Db 2334 TACCTCTGCTCTCCCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACCAGAGCGGAAG 2393  
Qy 421 CAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAATATAGAAATATTCTCTCAACAGCCTGT 480  
Db 2394 CAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAATATAGAAATATTCTCTCAACAGCCTGT 2453  
Qy 481 CCAGGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTCTGCTCAGCATCGACAGCC 540  
Db 2454 CCAGGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTCTGCTCAGCATCGACAGCC 2513  
Qy 541 TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGGACA 600  
Db 2514 TCTCCCTGTGCGATGACAGTGAAGAACCCATCAAAAGGCTCTCAAAAGCTCGCGGGACA 2573  
Qy 601 CCCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTTCCAGCTTGA 660  
Db 2574 CCCTGAGCTCAGCGGTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTTCCAGCTTGA 2633  
Qy 661 ACATGTGTGTCGCGCGGGCGGCCCAACGACACCCCAAGCTATTTCATATGTTGTAAG 720  
Db 2634 ACATGTGTGTCGCGCGGGCGGCCCAACGACACCCCAAGCTATTTCATATGTTGTAAG 2693  
Qy 721 TCCAAATACAGTCTCTTAATGTGAACACCTGACACATCCGGGAGTTTCCACCGGTCAAAAG 780  
Db 2694 TCCAAATACAGTCTCTTAATGTGAACACCTGACACATCCGGGAGTTTCCACCGGTCAAAAG 2753  
Qy 781 TGGGAGACATCGCACCTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCCAGCTTGTGTA 840  
Db 2754 TGGGAGACATCGCACCTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCCAGCTTGTGTA 2813  
Qy 841 CCAAAGACGGGAGCGCTTTCGCTACGACATGAGAGTGCAGAGTCCGGGATCGACCTGC 900  
Db 2814 CCAAAGACGGGAGCGCTTTCGCTACGACATGAGAGTGCAGAGTCCGGGATCGACCTGC 2873  
Qy 901 AGTGACACTGGCCCTGTATGGCAGCTTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGC 960  
Db 2874 AGTGACACTGGCCCTGTATGGCAGCTTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGC 2933  
Qy 961 TGGAGAACAGGGCCCTAACCTGCGCTCCACGCGGGCTCCACACTGCGCGAAGACAGCCT 1020  
Db 2934 TGGAGAACAGGGCCCTAACCTGCGCTCCACGCGGGCTCCACACTGCGCGG-AAGCAGCCT 2992  
Qy 1021 TCCTGCTCGGTGACATGCTGCCCTGAAACACAGAGCTCAGCGCTTCCAGGGGATGTTG 1080  
Db 2993 TCCTGCTCGGTGACATGCTGCCCTG-AAACAGAGCTCAGCGCTTCCAGGG--ATC 3048  
Qy 1081 NCCAGCCCCCGGCTCARCAGNTGGGAAACAGGGCCCTCGNCAGCAGCAGTNGGGGG 1140  
Db 3049 TGCCAGCCCCCGGCTCAGCAGTGGGACCCAGGGCTCGCAGCAGCAAG-----GTGGGG 3101  
Qy 1141 CAAGCNAGATGCTCCAGGATTTTCANCTGAGCCCTGCCANCCCTGCTGAADA 1200  
Db 3102 GCAAGCAGAAATGCTCCAGGATTTTCACAC---TGAGCCCTTGCCTCCAGCTGCTGAGAA 3158  
Qy 1201 ARAAYTNCGCCACGCTGAAGACAGAGAGGATGNCAGAGTTNNACCTTGGGGAA 1260  
Db 3159 AACACT--CCGCCACGTGAAGAGACA---GAGAGGATGGCAGGAGTTACCTCGGGAAA 3212  
Qy 1261 AAAAAACAGGGATCTTTNTTCTGCCCTGCTCCAGTNCAGAGTTGGCCTGNAACCCGCTGG 1320  
Db 3213 CAAACAGGATCTTCTGCTCCCTGCTCCAG-----TCCAGTTGGCCTGACCCGCTGG- 3264  
Qy 1321 ANTCACTGACATTTGTTGGCAGACAGGGGAGCAGAGCTTCCAGCTTGGGTCAAGAGGG 1380  
Db 3265 -ATCAGTGACATTTGCTGGCAGA-CAGGGGAGAGCAGCTTCCAGCTTGGGTCAAGAGGG 3322  
Qy 1381 GTGGGAGAGCCCTTCGCGCCCTCACCTNCCAGGCTGCTGNAGAGTGTCAAGTGTGA 1440  
Db 3323 GTGGGAGAGCCCTTCGCGCCCTCACCT-CCAGGCTGCTGTG-AGAGTGTCAAGTGTGA 3380  
Qy 1441 AGGNCNCCAAANTCTCAGNTTTCAGTGCAGAACACAGGTCNAGCAGGATATGCCCGCCGNTA 1500

Db 3381 AGGGCCCCAACTC---AGGTCAGTGCAGAACCCAGGT-CAGCAGGTATGCCGCCCGCTAG 3436  
Qy |||||  
Db 1501 GGTTAANNNGGGGGCCCTCTNAAAACCCCTTGCTNGGCCCTNCACCTNGGCCAGCTCANCCC 1560  
|||  
Db 3437 GTTAA---GGGGGCCCTTAACACCCCTTGCTGSCCTCA---CCTGGCCAGCTCA-CCC 3488  
|||  
Qy 1561 CTTTGGGTGTAGGGGAAAGAAATCCCTGAGACCCCTGGGAAGGTWCCCTGGTAGAATACAC 1620  
|||  
Db 3489 CTTTGGGTGTAGGGGAAAGAAATCCCTGAGACCCCTGGGAAGGT-CCCTGGTAGAATACAC 3547  
|||  
Qy 1621 CACACTTTTCAGTTGTTGCAACACAGGTCCGTGAGTTCACCTCTGGTTCCAGCAGGACC 1680  
|||  
Db 3548 CACACTTTTCAGTTGTTGCAACACAGGTCCGTGAGTTCACCTCTGGTTCCAGCAGGACC 3607  
|||  
Qy 1681 AAGAAGGTGTGAAGTGAAGTGGTTCCTCAGTNCACCAGACATGTGCCCTTTGCTGCTG 1740  
|||  
Db 3608 AAGAAGGTGTGAAGTGAAGTGGTTCCTCAGT-CCCCAGACATGTGCCCTTTGCTGCTG 3666  
|||  
Qy 1741 GCTACCACTCTTCCCGAGCAGCAGGCCCGAGCCCTTCAGGCCCGCAGCACTGCCCCAG 1800  
|||  
Db 3667 GCTACCACTCTTCCCGAGCAGCAGGCCCGAGCCCTTCAGGCCCGCAGCACTGCCCCAG 3726  
|||  
Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTCTAAAGTGAAGGTGATGCAAGGATATGCCCTGA 1860  
|||  
Db 3727 ACTCGCTGGCACTCAGTTCCCTCATCTCTAAAGTGAAGGTGATGCAAGGATATGCCCTGA 3786  
|||  
Qy 1861 CAGGAACAGTCTGTGGATGGACATCATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTG 1920  
|||  
Db 3787 CAGGAACAGTCTGTGGATGGACATCATCAGTGTCTAA--GGAAAGCAGCAGAGAGAGAGCGC 3844  
|||  
Qy 1921 TCCGGCGCCCGCAGNCCCACTNATCAGTGTNCCAGCGTGTCTGCTGCTGCTGCTGCTGCTG 1980  
|||  
Db 3845 TCCGGCGCCCGCAGCAGCCCACTATCAGTGT--CCAGCGTGTCTGCTGCTGCTGCTGCTG 3893  
|||  
Qy 1981 GCTNCAGNATCANCACTGACACTNCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
|||  
Db 3894 GCACAGCTCAGCATCACATGACACTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3951  
|||  
Qy 2041 CCGNACGCACTTTCACNTCTGATGNACTTCAAGCACTTTCATGCTGCTGCTGCTGCTGCTG 2100  
|||  
Db 3952 CCG-ACGGCACTTTGCGAC-TCGATG-ACCTCAAGCACTTTCATGCTGCTGCTGCTGCTG 4008  
|||  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGGNCAGTANGCAANGCCAGGAGATGG 2160  
|||  
Db 4009 GGGCAGGCVAGGCGAGTGACACTGT-----AGGAGCATAGCAAGCCAGGAGATGG 4058  
|||  
Qy 2161 GGTGNAAGGNCACAGCTTGTAGCTGTCCANCATGATGTGACATNCCCTCAAACTCTTN 2220  
|||  
Db 4059 GGTGAA--GGGACACAGCTTTGAGCTGTCCA-CATGCACTGTGACT-CCTCAAACTCTTT- 4113  
|||  
Qy 2221 NCCAGNATTTCTCTAAGAAATAGCANCCTTNCCTTCCAGCTTCCAGCTTACGCTCTTCT 2280  
|||  
Db 4114 --CGAGATTTCTTAAGAAATAGCACCCCTT--TTCCCCATTTGCCCGAGCTTAGCTCTTCT 4169  
|||  
Qy 2281 CCCAGGGAGCTANTCTCAGGACTCACGTAGCATTTAAATCAGGTGTGNAATCGTCAGGGGG 2340  
|||  
Db 4170 CCCAGGGAGCTA-CTCAGGACTCAGGTAGCATTTAAATCAGCTGTG-AATGTCAGGGGG 4227  
|||  
Qy 2341 TGTCTGTAGCTTCAACCTCTCGGGGAGGGGAGCCCGAGACTCCGTGGGAGAGCTCAT 2400  
|||  
Db 4228 TGTCTGTAGCTTCAACCTCTCGGGGAGGGGAGCCCGAGACTCCRTGGGAGAGCTCAT 4287  
|||  
Qy 2401 TCCACATCTTCCAGAGCAGCTTTTNGTCCAGCTGTCCACATTTGAGTCAGCTGCTCCC 2460  
|||  
Db 4288 TCCACATCTTCCAGAGCAGCTTTT--GTCCAGCTRTCCACATTTGAGTCAGCTGCTCCC 4346  
|||  
Qy 2461 GGGGAGAGAGCCCGCCCGCCAGCACATAAAGAACTGCAAGCTTGGTACTGCAAGTCTG 2520  
|||  
Db 4347 GGGGAGAGAGCCCGCCCGCCAGCACATAAAGAACTGCAAGCTTGGTACTGCAAGTCTG 4406  
|||  
Qy 2521 GGTGTAGAGAACTTTTGTAAAGCAATAAAGTTTGGGTGTATGACAAATGTT 2572  
|||  
Db 4407 GGTGTAGAGAACTTTTGTAAAGCAATAAAGTTTGGGTGTATGACAAATGTT 4458  
|||

RESULT 5  
US-60-659-397-566  
; Sequence 566, Application US/60659397  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: CHANG, Sheng-Yung  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001470  
; CURRENT APPLICATION NUMBER: US/60/659,397  
; CURRENT FILING DATE: 2005-03-09  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 566  
; LENGTH: 4684  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-659-397-566  
  
Query Match 67.7%; Score 1781.4; DB 15; Length 4684;  
Best Local Similarity 89.1%; Pred.No. 0;  
Matches 2292; Conservative 10; Mismatches 183; Indels 87; Gaps 35;  
  
Qy 1 CCCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 60  
Db 2150 CCCCTCTCACAGCCCGAGCCATCCAGAGGGGCTGAGGAAGAGCCCATCCACCGCGTGT 2209  
  
Qy 61 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 120  
Db 2210 CTGAGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGA 2269  
  
Qy 121 GCCCTTGGAGGGGAGAAATATAAGAAACAAGACATCAACCGCCAAATCAAGCCAAATTACC 180  
Db 2270 GCCCTTGGAGGGGAGAAATATAAGAAACAAGACATCAACCGCCAAATCAAGCCAAATTACC 2329  
  
Qy 181 ACCAGACCTTCATGCTCCAGCCGAGAGAGCTTTCGCAAGGGCCCGAGGGCCCGGCGAG 240  
Db 2330 ACCAGACCTTCATGCTCCAGCCGAGAGAGCTTTCGCAAGGGCCCGAGGGCCCGGCGAG 2389  
  
Qy 241 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTTCCAGCTCTCTCCACAGAGCCCGCAG 300  
Db 2390 CTGAGGAGACAAACAGGAGAGCCCTTAAGCTTCCAGCTCTCTCCACAGAGCCCGCAG 2449  
  
Qy 301 AGCCAAACAAGTCTCTCTCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGAAACCTT 360  
Db 2450 AGCCAAACAAGTCTCTCTCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGAAACCTT 2509  
  
Qy 361 TACCTCTGTCTCTCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCCAGAGCGGAAAG 420  
Db 2510 TACCTCTGTCTCTCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCCAGAGCGGAAAG 2569  
  
Qy 421 CAACCGCTCCCGAGCAGCAACTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCTGT 480  
Db 2570 CAACCGCTCCCGAGCAGCAACTGACAGAGCTGGAAATAGAAATATTCTCTCAACAGCTGT 2629  
  
Qy 481 CCCAGCAATTTTCTCTGGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATGCAAGCC 540  
Db 2630 CCCAGCAATTTTCTCTGGAGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATGCAAGCC 2689  
  
Qy 541 TCTCTCTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGAGCA 600  
Db 2690 TCTCTCTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGAGCA 2749  
  
Qy 601 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCGGAGGCTCGAAGCTCCAGCTGGA 660  
Db 2750 CCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCGGAGGCTCGAAGCTCCAGCTGGA 2809  
  
Qy 661 ACATGGTCTGGCCCGGGGCGGCCACCGACACCCCAAGCTATTTCATGCTGTGAAAG 720  
|||



Db 2810 ACATGGTGTGGCCGGGGGGGGCCACCCAGACACCCCAAGCTATTTCAATGGTGTGAAG 2869  
Qy 721 TCCAAATACAGTCTCTTAATAGTGAACAACCTGCAATCCCGGAGTTCCACCGGTCAAAG 780  
Db 2870 TCCAAATACAGTCTCTTAATAGTGAACAACCTGCAATCCCGGAGTTCCACCGGTCAAAG 2929  
Qy 781 TGGGAGACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGTCA 840  
Db 2930 TGGGAGACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGTSA 2989  
Qy 841 CCAAAGACGGGACGCTGTTGCTTACGATCAGATGGAGGTGCAGACTCCGGGATCGACCTGC 900  
Db 2990 CCAAAGACGGGACGCTGTTGCTTACGATCAGATGGAGGTGCAGACTCCGGGATCGACCTGC 3049  
Qy 901 AGTGCACACTGGCCCTGATGCGAGCTTGCCTGGAGCTGGAGGTCAAGCATGGCCAGC 960  
Db 3050 AGTGCACACTGGCCCTGATGCGAGCTTGCCTGGAGCTGGAGGTCAAGCATGGCCAGC 3109  
Qy 961 TGGAGAACAGCCCTAACCTGCTCCACCGCCGGCTCCACACTGCGGGAAGCAGCCT 1020  
Db 3110 TGGAGAACAGCCCTAACCTGCTCCACCGCCGGCTCCACACTGCGGGAAGCAGCCT 3168  
Qy 1021 TCCCTGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCCGTTTCCAGGGGATYTG 1080  
Db 3169 TCCCTGCTCGGTGACGATGCTGCCCTG- AAACACAGGCTCAGCCGTTTMCAGGG- ---ATC 3224  
Qy 1081 NCCAGCCCCCGGCTCARGAGTGGGAACACGAGGCTCGNACAGNAGNAGTNGGGG 1140  
Db 3225 TGCAGCCCCCGCTCAGCAGTGGGACGAGGCTCGCAGCAGCAAG- -----GTGGG 3277  
Qy 1141 CAGCNAGATGCTCCAGGATTTCAACCTGAGCCGNTGCCCNACCTGCTGAADA 1200  
Db 3278 GCAAGCAGATGCTCCAGGATTTCAAC- ---TGAGCCCTGCCCCACCTGCTGAGAA 3334  
Qy 1201 AACAAYTNCGCCAGCTGAGAGACAGAGGATGNCAGGATTTNACCTYGGGAA 1260  
Db 3335 AACACT- ---CGGCCAGTGAAGAGACA- ---GAGGAGATGCGAGGATTAACCTCGGAAA 3388  
Qy 1261 ACAAACAGGAGATCTTTTTCGCCCCCTGCTCCAGTNCAGTGGCTGAGCCGCTGAG 1320  
Db 3389 CAACAGGATCTTCTGCCCCCTGCTCAG- -----TGAGTTGGCTGACCCGCTTG- 3440  
Qy 1321 ANTCAGTGACATTTGTCAGANACAGGGAGAGCTTCCAGCTCGGTGAGAAGG 1380  
Db 3441 -ATCAGTGACATTTGCTGGCAGA- CAGGGAGAGCAGCTTCCAGCTCGGTGAGAAGG 3498  
Qy 1381 GTGGCCAGCCCTTCGGCCCTCACCCTNCAGGCTGCTGAGAGTGTCAAGTGTGA 1440  
Db 3499 GTGGCCAGCCCTTCGGCCCTCACCCT- CCAGGCTGCTGTG- AGAGTGTCAAGTGTGA 3556  
Qy 1441 AGGNCCTCAANCTCAGNTTTCAGTGACAGAACACAGGTTNACAGGATGCGCCGCGNTA 1500  
Db 3557 AGGCCCAAACTC- ---AGTTTCAGTGAGAACCAAGGT- CAGCAGGTATGCCGCCCTAG 3612  
Qy 1501 GGTTAANNNGGGGCCCTCTNAACCCCTTGCTNGGCTNCACCTNGGCCAGCTCANCCC 1560  
Db 3613 GTTAA- ---GGGGCCCTCTAAACCCCTTGCTGCGCTCA- ---CCTGGCCAGCTCA- CCC 3664  
Qy 1561 CTTTTGGGTAGGGGAAAAGATTCCTGACCTCGGGAAGGCTWCCCTGGTAGAATACAC 1620  
Db 3665 CTTTTGGGTAGGGGAAAAGATTCCTGACCTCGGGAAGGCT- ---CCTGGTAGAATACAC 3723  
Qy 1621 CACATTTTCAGGTGTGCAACAGGCTGAGTTGACCTCTGTTTCAGCTGACCAAGACC 1680  
Db 3724 CACATTTTCAGGTGTGCAACAGGCTGAGTTGACCTCTGTTTCAGCTGACCAAGACC 3783  
Qy 1681 AAGAAGGTGTGAAGTGAAGTGTCTCAGTNCACAGATGTCGCCCTTTGCTGCTG 1740  
Db 3784 AAGAAGGTGTGAAGTGAAGTGTCTCAGT- ---CCACAGATGTCGCCCTTTGCTGCTG 3842  
Qy 1741 GCTACCACTCTTCCCCAGAGCAGCAGGCCCGGAGCCCTTTAGGCCAGCACTGCCCCAG 1800  
Db 3843 GCTACCACTCTTCCCCAGAGCAGCAGGCCCGGAGCCCTTTAGGCCAGCACTGCCCCAG 3902

Qy 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGCGAGTATGCCTGA 1860  
Db 3903 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGCGAGTATGCCTGA 3962  
Qy 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAGNAAAGCAGCAGAGAGAGCY 1920  
Db 3963 CAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAA- ---GGAAAGCAGCAGAGAGAGC 4020  
Qy 1921 TCCGGGGCCCCCAGNCCCCTACTNATCAGTGTNCCAGCGTGTNGGTTNCCCAGNAGACA 1980  
Db 4021 TCCGGGGCCCCCAGNCCCCTACTNATCAGTGT- ---CCAGCGTGTGGTTCCC- ---CAGA 4069  
Qy 1981 GCTNACGNATCANCACTGACATNACACCTTNGCCCTGCCCCCTNGGCCANGAGGTTACTG 2040  
Db 4070 GCACAGCTCAGCATCACTGACACTCACCTGCCCCCTGCCCCCTGG- ---CCAGAGGTTACTG 4127  
Qy 2041 CCGNAGCGCACTTTCGACNTCTCATGNACCTCAAGACATTTTCATGGCTNGCCCTCTNNG 2100  
Db 4128 CCG- ---ACGGCACTTTGAC- ---TCTGATG- ---ACCTCAAGACATTTTCATGGCTGCCCTCTGGCA 4184  
Qy 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGNAGCATANGCAANGCCAGGAGATGG 2160  
Db 4185 GGCAGGGYAGGCGAGTGACACTGT- -----AGGACATAGCAAGCCAGGAGATGG 4234  
Qy 2161 GGTGNAAGGANCACAGTCTTGAGCTGTCCANCATGTCATGTGACTCTNCTCAAACTCTTN 2220  
Db 4235 GGTGAA- ---GGGACACAGTCTTGAGCTGTCCA- ---CATGCATGTGACT- ---CCTCAAACTCTT- 4289  
Qy 2221 NCCAGNATTTCTTAAGATAGCANCCCCCTTNGCCCTTNGCCCTGAGCTTACCTCTTCT 2280  
Db 4290 - ---CCAGATTTTCTTAAGATAGCACCCCC- ---TTCCCCATTTGCCCCAGCTTAGCCTCTTCT 4345  
Qy 2281 CCAGGGGAGCTANCTCAGGACTCAGTAGCTTAATATCAGTGTGNAATCGTCAGGGG 2340  
Db 4346 CCAGGGGAGCTA- ---CTCAGGACTCAGTAGCTTAATATCAGTGTG- ---AATCGTCAGGGG 4403  
Qy 2341 TGTCTGTAGCTCAACCTCTCTGGGCGAGGGAGCGCGAGACTCCGTGGGAGAGCTCAT 2400  
Db 4404 TGTCTGTAGCTCAACCTCTCTGGGCGAGGGAGCGCGAGACTCCRTGGGAGAGCTCAT 4463  
Qy 2401 TCCACATCTTGCAGAGAGAGCTTTNGTCCAGCTGTCCACATTCAGTTCAGCTGCTCC 2460  
Db 4464 TCCACATCTTGCAGAGAGAGCTTT- ---GTCCAGCTRTCCACATTCAGTTCAGCTGCTCC 4522  
Qy 2461 GGGGAGAGAGCCCCCGCCCCCAGACACATAAAGAACTGCAGCCTTGTGTAAGTGTG 2520  
Db 4523 GGGGAGAGAGCCCCCGCCCCCAGACACATAAAGAACTGCAGCCTTGTGTAAGTGTG 4582  
Qy 2521 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTGACAAATGTT 2572  
Db 4583 GGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGTGACAAATGTT 4634

## RESULT 6

US-60-659-397-12129  
; Sequence 12129, Application US/60659397  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: CHANG, Sheng-Yung  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001470  
; CURRENT APPLICATION NUMBER: US/60/659,397  
; NUMBER OF SEQ ID NOS: 47859  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12129  
; LENGTH: 65967  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-60-659-397-12129

Query Match 36.9%; Score 971.8; DB 15; Length 65967;  
Best Local Similarity 84.2%; Pred. No. 1.6e-222; Mismatches 183; Indels 87; Gaps 35;  
Matches 1483; Conservative 9;  
QY 811 AGATCCCAAGCTCAGCCCTTACGTTGGTCAACCAAGAGGGGAGGCTGTTCGTACGACA 870  
DB 58244 AGATCCCAAGCTCAGCCCTTACGTTGGTISACCAAGAGGGGAGGCTGTTCGTACGACA 58303  
QY 871 TGGAGGTGCAGACTCGGGCATCGACTGCGATGACACACTGCGCCCTCTGATGGCAGCTTCG 930  
DB 58304 TGGAGGTGCAGACTCGGGCATCGACTGCGATGACACACTGCGCCCTCTGATGGCAGCTTCG 58363  
QY 931 CCGTGAGCTGGAAGGTCGAAGCATGCCAGCTGGAGAACAGGCCCTAAACCTGCCCTCCAC 990  
DB 58364 CCGTGAGCTGGAAGGTCGAAGCATGCCAGCTGGAGAACAGGCCCTAAACCTGCCCTCCAC 58423  
QY 991 CGCCGGCTCCACACTGCCGGAAAGCAGCCCTTCTGCTCGGTGCACGATGCTGCCCTGAAA 1050  
DB 58424 CGCCGGCTCCACACTGCCGG--AAGCAGCCTTCTGCTCGGTGCACGATGCTGCCCTG-AA 58481  
QY 1051 ACACAGGTTCAGCCGTTCCAGGGGATYTGNCAGCCGCCCGGCTCARGAGNTGGGAACC 1110  
DB 58482 ACACAGGTTCAGCCGTTCCAGGGG---ATCTGCCAGAGCCCGGCTCAGCAGTGGGACCCAG 58538  
QY 1111 AGGGCTCGNAGAGNAGNAAGTNGGGGCAAGCNAGNATGCCCTCCAGGATTTTCACAN 1170  
DB 58539 GGCCTCGCAGCAGCAAG-----GTGGGGGCAAGCAGAAATGCCCTCCAGGATTTTCACAC 58591  
QY 1171 CCGTAGCCCTNCCCTCCCTGCTGAADAAAACAYTNCGCCGACGCTGAAGAGACAGAAG 1230  
DB 58592 C---TGAGCCCTGCCCCCTGCTGAGAAACACT--CGGCCACGCTGAAGAGACA---- 58642  
QY 1231 GAGGATGNCAGGAGTTNNACCTYGGGGAACAAACAGGGATCTTTNTTCGCCCTTCG 1290  
DB 58643 GAGGAGGATGGCAGGAGTTACCTCGGGAACAAACAGGATCTTCTCTGCTGCTGCCAG- 58701  
QY 1291 TCCAGTNCAGTTGGCTGNACCCGCTTGGANTCAGTACGACCATTTGTTGGCAGANCAGG 1350  
DB 58702 -----TCGAGTTGGCCCTGA---CCGCTTGGATCAGTACGACCATTTGCTGGCAGA--CAGGG 58752  
QY 1351 GAGAGCAGCTTCAGCCTCGGTCAGAAAGGGTGGGCGAGCCCTTCGGGCCCTCACCTTNC 1410  
DB 58753 GAGAGCAGCTTCAGCCTCGGTCAGAAAGGGTGGGCGAGCCCTTCGGGCCCTCACCTT-C 58811  
QY 1411 CAGGCTGCTGTGACAGTGTCAAGTGTGAAGGNCACCAANCTCAGGNTTCAGTGCAGA 1470  
DB 58812 CAGGCTGCTGTG-AGAGTGTCAAGTGTGAAGGGCCCAAACTC---AGGTTTCAGTGCAGA 58867  
QY 1471 ACCAGGTNCAGCAGGTATGCCCGCCGNTAGGTTAANNNGGGGCCCTCTNAAACCCCTTG 1530  
DB 58868 ACCAGGT-CAGCAGGTATGCCCGCCGNTAGGTTAA-----GGGGGCCCTCTAAACCCCTTG 58922  
QY 1531 CCGTNGCCCTNCACTTNGGCCACTCANCCCTTTTGGGTGTAGGGGAAGAAATCCCTGA 1590  
DB 58923 CCGTNGCCCTCA---CCTGGCCAGCTCA-CGCCCTTTGGGTGTAGGGGAAGAAATCCCTGA 58978  
QY 1591 CCCTGGGAAGGCTWCCCTGGTGAAGATACACACACTTTTCAGGTTGGTGCACACAGGTC 1650  
DB 58979 CCCTGGGAAGGCT-CCCTGGTGAAGATACACACACTTTTCAGGTTGGTGCACACAGGTC 59037  
QY 1651 CTGAGTTGACCTCTGGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCA 1710  
DB 59038 CTGAGTTGACCTCTGGTTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCA 59097  
QY 1711 GTNCCCGACACATGTGCCCTTTGCTGCTGGCTACCACTCTTCCCGCAGACGAGGCC 1770  
DB 59098 GT-CCCCAGACATGTGCCCTTTGCTGCTGGCTACCACTCTTCCCGCAGACGAGGCC 59156  
QY 1771 CGAGGCCCTTCAGGCCCGCAGCACTGCCCGCAGTCCGCTGGCACTCAGTTCCCTCATCTGA 1830  
DB 59157 CGAGGCCCTTCAGGCCCGCAGCACTGCCCGCAGTCCGCTGGCACTCAGTTCCCTCATCTGA 59216

QY 1831 AAGGTGAAGGCTGATGACGATATGCTCTGACAGGAACAGTCTGTGGATGGACATGATCAG 1890  
DB 59217 AAGGTGAAGGCTGATGACGATATGCTCTGACAGGAACAGTCTGTGGATGGACATGATCAG 59276  
QY 1891 TGCTNAAGNAAAGCAGCAGAGAGAGAGCGYTCGCGCGCCCCAGNCCCCCCTNATCAGTGT 1950  
DB 59277 TGCTAA--GGAAAGCAGCAGAGAGAGAGCGTCCGCGCCCCCAGGCCCTATCAGTGT-- 59332  
QY 1951 NCCAGCGTGTCTNGGTTNCCCGCAGNAGCAGACTNCAAGCATCAGCTGACATNCACCC 2010  
DB 59333 --CAGCGTGTCTGCTCC-----CAGAGCAGCTCAGCATCAGCTGACACTCACC 59383  
QY 2011 TNGCCCTGCCCTTNGGCCANGAGGCTACTCGGCGNAGCGGCACTTTTGACNTCTGTAGNACC 2070  
DB 59384 CTGCGCTGCCCTCG--CCAGAGGCTACTCGG-ACGCGACTTTTGAC-TCGTATG-ACC 59438  
QY 2071 TCAAAGCACTTTTCATGCTNGCCCTCTNNCGCAGGNCAGGNCAGGNCAGTGCANCT 2130  
DB 59439 TCAAAGCACTTTTCATGCTGCTCCCTCTGCGAGGCGAGGCGAGGCTGACACT----- 59491  
QY 2131 GTAGNAGCATANGCAANGCCAGGAGATGGGGTGNPAGGGGANCACAGTCTTTGAGCTGTCC 2190  
DB 59492 ---GTAGGAGCATAGCAAGCCAGGAGATGGGGTGAA--GGGACACAGTCTTTGAGCTGTCC 59546  
QY 2191 ANCATGCAATGTGACTNCCCTCAAACCTCTTNNCCAGNATTTCTTAAGAAATAGCANCCCC 2250  
DB 59547 A-CATGCAATGTGACT-CCTCAAACCTCTT---CCAGATTTCTTAAGAAATAGCACCCCC- 59600  
QY 2251 TTNCCCCATTGCCCAGCTTACGCTCTTCTCCAGGGGAGCTANCTCAGGACTCACGTAG 2310  
DB 59601 -TTCCCATTTGCCCAGCTTACGCTCTTCTCCAGGGGAGCTA-CTCAGGACTCACGTAG 59658  
QY 2311 CATTAATCAGCTGTGNATCGTCAGGGGTGTCTGTAGCCTCAACCTCTGGGGCAGG 2370  
DB 59659 CATTAATCAGCTGTG--AATCGTCAGGGGTGTCTGTAGCCTCAACCTCTGGGGCAGG 59717  
QY 2371 GGAGCCGAGACTCCGTGGGGAAGAGTCAATTCACACATCTTTCGCAAGACAGCTTTTNGTC 2430  
DB 59718 GGAGCCGAGACTCCGTGGGGAAGAGTCAATTCACACATCTTTCGCAAGACAGCTTTT- 59776  
QY 2431 CAGCTGTCCAATTCAGTTCAGACTGTCTCCCGGGGAGAGAGCCCGGCCCCCGACACATAA 2490  
DB 59777 CAGCTGTCCAATTCAGTTCAGACTGTCTCCCGGGGAGAGAGCCCGGCCCCCGACACATAA 59836  
QY 2491 AGAACTGCGACCTTGGTACTCGACAGTCTGGGTGTGAGAGAACTTTTGTAGCAATAA 2550  
DB 59837 AGAACTGCGACCTTGGTACTCGACAGTCTGGGTGTGAGAGAACTTTTGTAGCAATAA 59896  
QY 2551 GTTTGGGTGATGACAAATGTT 2572  
DB 59897 GTTTGGGTGATGACAAATGTT 59918

RESULT 7

US-11-060-756-560  
; Sequence 560: Application US/11060756  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William Martin  
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug  
; TITLE OF INVENTION: Target Genes  
; FILE REFERENCE: AM101083 (031896-042000)  
; CURRENT APPLICATION NUMBER: US/11/060,756  
; CURRENT FILING DATE: 2005-02-18  
; NUMBER OF SEQ ID NOS: 303284  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 560  
; LENGTH: 1400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-060-756-560









[illegible][illegible]

GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary  
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 32424  
; LENGTH: 3658  
; TYPE: DNA  
; ORGANISM: Macaca Mulatta  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(3658)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-680-473-32424

Query Match 11.2%; Score 295.6; DB 15; Length 3658;  
Best Local Similarity 28.7%; Pred. No. 1.4e-60;  
Matches 645; Conservative 1; Mismatches 1491; Indels 107; Gaps 15;  
QY 329 GAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTACCTCTGTCTCTCTCTGGAGCCGCC 388  
DB 1522 GAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTACCTCTGTCTCTCTCTGGAGCCGCC 1581  
QY 389 TGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACCGTCCCGAGCAGGAATGTCAGCA 448  
DB 1582 TGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACCTTCCCGAGCAGGAATGTCAGCA 1641  
QY 449 GCTGGAATAGAAATATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGA 508  
DB 1642 GCTGGAATA----- 1651  
QY 509 GCAAAATCTCTGTGCTCAGCATCGACAGCCTCTCCCTGTGGATGACAGTGAGAAAGAA 568  
DB 1652 -NN 1710  
QY 569 CCATCAAGAGCCTCTCAAGCTCGCGGACACCTCAGCTCAGCGCTACACTCTGGAG 628  
DB 1711 NNN 1770  
QY 629 CAGCAGGCGGAGCTCGAAGCTCCAGCTGGAACATGGTGTGCGCCCGGGCGGCCAC 688  
DB 1771 NNN 1830  
QY 689 CGACACCCCAAGTATTTCATGGTGTGAAGTCAAAATACAGTCTCTTAATGGTGAACA 748  
DB 1831 NNN 1890  
QY 749 CTTGCACATCGGAGGTTCCACCGGTCAAAGTGGGAGACATCGCCACTGGCATCAGCAG 808  
DB 1891 NNN 1950  
QY 809 CCAGATCCAGCTGACGCTTACAGCTTGGTCAACCAAGACGGGAGCGCTGTTCGCTACGA 868  
DB 1951 NNN 2010  
QY 869 CATGGAGGTGCCAGACTCGGGCATTCGACTGTGACATCGACCTGACACTGGCCCTGATGGCAGCTT 928  
DB 2011 NNN 2070  
QY 929 CGCTGGAGGTGAGGCTCAAGCATGGCCAGCTGGAGAAACAGGCCCTAACCCCTGCCCTCC 988  
DB 2071 NNN 2130

QY 989 ACCGCCGGCTCCACACTGCCGAAAGCAGCGCTTCTCTGCTCGGTGCACGATGCTGCCCTGA 1048  
DB 2131 NNN 2190  
QY 1049 AAACACAGGCTCAGCGCTTCCAGGGGATTGNCACAGCCCCCGGCTCARCAGTGGAA 1108  
DB 2191 NNN 2250  
QY 1109 CCAGGGCTCGCAGCAGCAGCAGGTTGGGGCAAGCAGAAATCGCTCCAGGATTTTCAC 1168  
DB 2251 NNN 2310  
QY 1169 ANCTGAGCCCTGCCCAACCTCTGTGAADAAACAYTTCGCCCACTGTAAGACAGACA 1228  
DB 2311 NNN 2370  
QY 1229 AGGAGGATGNCAGGAGTTTNACTTGGGGAAACAAACAGGGATCTTTTTCCTGCCCC 1288  
DB 2371 NNN 2430  
QY 1289 GCTCCAGTNCAGTTGGCCTGNAACCCGCTTGGGANTCAGTGACCATTTGTTGGCAGAC 1348  
DB 2431 NNN 2490  
QY 1349 GGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCC 1408  
DB 2491 NNN 2550  
QY 1409 NCAGGCTGCTGTGAGAGTGTCAAGTGTGAAGGCCCAAACTCAGGTTTCAGTTGCA 1468  
DB 2551 NNN 2610  
QY 1469 GAACGAGTNCAGCAGGATGCGCCGCCGCTAGTTAANGGGGGCCCTCTNAAACCCCT 1528  
DB 2611 NNN 2670  
QY 1529 TGCTTGGCCTNCACCTTNGGCCAGCTCANCCTCTTTTGGGTGTAGGGGAAAGAAATGC 1588  
DB 2671 NNN 2730  
QY 1589 GACCTGGGAAGGCTWCCCTGGTAGAATAACACACATTTTCAGGTTGTGCAACAGG 1648  
DB 2731 NNN 2790  
QY 1649 TCCTGAGTTGACCTCTGGTTTCAGCAGGACCAAGAGGTGTCTAAGTGAAGTGTCT 1708  
DB 2791 NNN 2850  
QY 1709 CAGTNCACAGACATGTGCCCTTTGCTGTGCTACCACTTCTCCCGCAGCAGCAGGC 1768  
DB 2851 NNN 2910  
QY 1769 CCCGAGCCCTTCAGGCCAGCAGCTGCCCCAGACATCGCTGGCAGCTCAGTCCCTCATCTG 1828  
DB 2911 NNN 2970  
QY 1829 TAAAGTGAAGGTGATGCAGGATATGCTGTGACAGGACAGTCTGTGTGATGGACATGC 1888  
DB 2971 NNN 3030  
QY 1889 AGTGTAAAGGAAAGCAGCAGAGAGAGAGCGTCCGCGGCCCGCAGNCCCACTNATCAGT 1948  
DB 3031 NNN 3084  
QY 1949 GTNCCAGCGTGTGGTTTCCCGCAGCAGCAGCTCAGCTCAGNCACTCAGCAGTGCAC 2008  
DB 3085 CAGCGTCCAGCGTGGCGGCCCGCCAGAGCAGCT-----CACAAACACTCA 3133  
QY 2009 CTTNGCCCTCGCCTGGCCAGAGGCTACTGCGNACGAGCAGCTTTCGACNTCTGATGNA 2068  
DB 3134 CCCTGCCCTGCCCT--GGGCGGAGGGTACTGCTG-ACAGCACTTTGCAC-TCTGATG-A 3188

Qy	2069	CCTCAAGCATTTCATGGCTGCTCTNNGGCAGGNCAGGNCAGGNCAGTGCAN	2128
Db	3189	CCTCAAGCATTTCATGGCTGCTCTNNGGCAGGNCAGGNCAGGNCAGTGCAGTG	3233
Qy	2129	CTGTAGGAGCATANGCAAGCCAGGAGATGGGTGNAAGGANCACAGTCTTCAGCTGT	2188
Db	3234	ACTCTGTAGGAGCATAGCAGTCCAGGAGATGGGTGNAAGG---ACAGTCTTCAGCTGT	3289
Qy	2189	CCAMCATCATGTGATCTNCCCTCAAACTCTTNNCCAGNATTTCTTAAGAATAGCAGCC	2248
Db	3290	CTA-CATGATGTGACTCTCAAACTCTTNNCCAGNATTTCTTAAGAATAGCAGCC	3342
Qy	2249	CCTNNCCCATGTCAGGCTAGCTCTTCTCCAGGAGGAGTANCTCAGGACTCAGT	2308
Db	3343	---TTCCTCATGTCAGGCTAGCTCTTCTCCAGGAGGAGCTA-CTCAGGACTCATGT	3398
Qy	2309	AGCATTAATCAGCTGTGNAATCGTCAAGGGGTGTCTGAGCTCTCAACTCTCTGGGCA	2368
Db	3399	AGCATTAATCAGCTGTGNAATCGTCAAGGGGTGTCTGAGCTCTCAACTCTCTGGGCA	3458
Qy	2369	GGGACGCGAGACTCCGTGGAGAGTCAATTCCTCAATCTTCGCAAGACAGCTTTNG	2428
Db	3459	GGGACGCGAGACTCCGTGGAGAGTCAATTCCTCAATCTTCGCAAGACAGCTTTNG	3516
Qy	2429	TCCAGCTGTCCATTTGAGTCAAGTCTCCGGGAGAGAGCCCGCCCGCCAGCACAT	2488
Db	3517	TCCAGCTGTCCATTTGAGTCAAGTCTCCGGGAGAGAGCCCGCCCGCCAGCACAT	3576
Qy	2489	AAAGAACTGCAGCTTGTGACTGCAGAGTCTGGGTGTGAGAGAACTTTTGAAGCAATA	2548
Db	3577	AAAGAACTGCAGCTTGTGACTGCAGAGTCTGGGTGTGAGAGAACTTTTGAAGCAATA	3634
Qy	2549	AAGTTTGGGTGATGACAAATGTT	2572
Db	3635	AAGTTTGGGTGATGACAAATGTT	3658
RESULT 13			
US-60-544-10785/c			
; Sequence 10785, Application US/60680544			
; GENERAL INFORMATION:			
; APPLICANT: Cooper, Matthew			
; APPLICANT: Kinch, Deborah			
; APPLICANT: Rosenberg, Michael			
; APPLICANT: Subramaniam, S. Sai			
; APPLICANT: Szak, Suzanne			
; APPLICANT: Li, Huo			
; APPLICANT: Bandaru, Raj			
; APPLICANT: Derbel, Maher			
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary			
; FILE REFERENCE: 2159029000			
; CURRENT FILING DATE: 2005-05-13			
; NUMBER OF SEQ ID NOS: 48714			
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0			
; SEQ ID NO 10785			
; LENGTH: 761			
; TYPE: DNA			
; ORGANISM: Macaca Fascicularis			
; FEATURE:			
; NAME/KEY: misc.feature			
; LOCATION: (1)...(761)			
; OTHER INFORMATION: n = A, T, C or G			
US-60-544-10785			
Query Match 9.8%; Score 257.4; DB 15; Length 761;			
Best Local Similarity 59.8%; Pred. No. 1.4e-51;			
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;			
Qy	1719	GACATGTGCCCCCTTGTGCTGCTACCACTCTTCCCCAGAGCAGAGCCCGGAGCCCC	1778
Db	761	GACACATGNNNNNGCTGTGCTATCANNNTCCCGAGCGCAGGNNNNNGAGCCCC	702

Qy	1779	TTCAGGCCAGCACTGCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAGGTGAA	1838
Db	701	TTCAGGCCAG--NN	644
Qy	1839	GGGTGATGAGGATATGCTTGCAGAGAACAGTCTGTGGATGGAATGATCATGCTTNAAG	1898
Db	643	NN	586
Qy	1899	GNAAGCAGCAGAGAGAGAGTCCGGCGCCCGCCAGNCCCACCTNATCAGGTGCCAGCGT	1958
Db	585	GGAAGCAGCAGAGAGAGAGTCCGGCGCCCGCCAGNCCCACCTNATCAGGTGCCAG	532
Qy	1959	GCTNGTTCNCCAGNAGCAGCTNCAAGNCACTGACACTNCACTNCACTNCACTNCACT	2018
Db	531	CGTCCGGGCCCGCCAGAGCAGCTCACGA-----ACACTCACCTG	488
Qy	2019	CCCTNCGCCANGAGGTACTCCGCGNACGGACCTTTGCACTTGTGATGATGATGATGAT	2078
Db	487	CCCGCGCCCGCGGAGGTATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNA	428
Qy	2079	CTTTCATGGCTNGCCCTCTTNNCGAGGNCAGGNCAGGNCAGGNCAGGNCAGGNCAGG	2138
Db	427	CTTTCATGGCTG---CCCTCTGCGAGGCGAGGTGAGTCTCTGTAGAGC-----	380
Qy	2139	CATANGCAAGCCAGGAGATGGGTGNAAGGGANCAAGTCTTTGAGCTGTCANCAATGCA	2198
Db	379	-----ANNNGCA	328
Qy	2199	TGTGACTNCTCAAACTCTTNNCCAGNATTTCTTAAGATAGCAGCCCTTNCCTCA	2258
Db	327	TGTGACTNCTCAAACTCTCT---CCAGATTTCTTAATAGCAACCAAC-----TTC	277
Qy	2259	TTGCCCGCAGCTTACCTCTTCTCCAGGGAGTCTCAGGACTCAGGACTTAAAT	2318
Db	276	TTGCCCGCAGCTTACCTCTTCTCCAGGGAGTCTCAGGACTCAGGACTTAAAT	218
Qy	2319	CAGCTGTGNAATCGTCAGGGGTGTCTCTAGCTCAACCTCTCTGGGCGAGGCGCG	2378
Db	217	CAGCTGTGAAATGTCAGGGGTGTCTCAGTACCTCAACCTCTCTGGGCGAGGCGCG	158
Qy	2379	AGACTCCGTGGAGAGCTCATTCACATCTTCCAGAGCAGCTTTTNGTCCAGCTGTC	2438
Db	157	AGACTCCGTGGAGAGCTCATTCACATCTTCCAGAGCAGCTTTTNGTCCAGCTGTC	99
Qy	2439	CACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGCCCGCCAGCAGCATAAAGAACTGC	2498
Db	98	CACAGCGAGCAGACTGCTCTCCGGGAGAGAGCCCGCCCGCCAGCAGCATAAAGAACTGC	39
Qy	2499	AGCCTTGTGCTAGTCTGGGTGTGAGAGAACTCTT	2537
Db	38	AGCGTTGTGCTAGTGC--AGTGTGGGTGTGAGAGAACTCTT	2
RESULT 14			
US-60-544-11056			
; Sequence 11056, Application US/60680544			
; GENERAL INFORMATION:			
; APPLICANT: Cooper, Matthew			
; APPLICANT: Kinch, Deborah			
; APPLICANT: Rosenberg, Michael			
; APPLICANT: Subramaniam, S. Sai			
; APPLICANT: Szak, Suzanne			
; APPLICANT: Li, Huo			
; APPLICANT: Bandaru, Raj			
; APPLICANT: Derbel, Maher			
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary			
; FILE REFERENCE: 2159029000			
; CURRENT FILING DATE: 2005-05-13			
; NUMBER OF SEQ ID NOS: 48714			
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0			
; SEQ ID NO 11056			
; LENGTH: 761			
; TYPE: DNA			
; ORGANISM: Macaca Fascicularis			
; FEATURE:			
; NAME/KEY: misc.feature			
; LOCATION: (1)...(761)			
; OTHER INFORMATION: n = A, T, C or G			
US-60-544-10785			
Query Match 9.8%; Score 257.4; DB 15; Length 761;			
Best Local Similarity 59.8%; Pred. No. 1.4e-51;			
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;			
Qy	1719	GACATGTGCCCCCTTGTGCTGCTACCACTCTTCCCCAGAGCAGAGCCCGGAGCCCC	1778
Db	761	GACACATGNNNNNGCTGTGCTATCANNNTCCCGAGCGCAGGNNNNNGAGCCCC	702

; SEQ ID NO 11056  
; LENGTH: 761  
; TYPE: DNA  
; ORGANISM: Macaca Fascicularis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(761)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-680-544-11056

Query Match 9.8%; Score 257.4; DB 15; Length 761;  
Best Local Similarity 59.8%; Pred. No. 1.4e-51;  
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;  
  
Qy 1719 GACATGTGCCCTTTGCTGCTGTACCACTCTTCCCGAGCAGCAGCGCCGAGCCCC 1778  
Db 1 GACACATGNNNNNGCTGCTGGCTATCANNNTCCCGAGCGCAGCAGNNNNNGAGCCCC 60  
  
Qy 1779 TTCAGGCCAGCACTGCCCGAGACTCGCTGGCACTCAGTTCCTCATCTCTAAAGGTGAA 1838  
Db 61 TTCAGGCCAG--NN 118  
  
Qy 1839 GGGTGATGCAGGATATCCCTGCAGGAACAGTCTGTGGATGGACATGATCAGTCTNAAG 1898  
Db 119 NNN 176  
  
Qy 1899 GNAAGCAGCAGAGAGAGCGYTCCGGCGGCCCGAGNCCCACTNATCAGTGTNCCAGCGT 1958  
Db 177 GGAAGCAGCAGCAGAGAGCGTCCGGCACCCAGCCCC-----ACTGTGAGGTCCAG 230  
  
Qy 1959 GCTNGGTTNCCCGAGNAGCAGCTNAGNATCANCACTGACACTNCACTTNGCCCTG 2018  
Db 231 CGTGGCGGCCCGCAGAGCAGCTCACGCA-----ACACTCACCTG 274  
  
Qy 2019 CCCTNGGCCANGAGGTACTGCCGACGCACTTTGCACNTCTGATGNACCTCAAAGCA 2078  
Db 275 CCGCGCCCGCGGAGGGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 334  
  
Qy 2079 CTTTCATGGCTNGCCCTCTNNGCAGGNCAGGNCAGTGCAGTGTAGNAG 2138  
Db 335 CTTTCATGGCTG---CCCTCTGGCAGGCGAGTGCAGTGTCTGTAGGAGC----- 382  
  
Qy 2139 CATANGCAAGCCAGGAGATGGGGTGNAAAGGGANCAAGTCTTGAGCTGTCCANCAATGCA 2198  
Db 383 -----ANNNGCA 434  
  
Qy 2199 TGTGACTNCTCAAACCTCTTNNCCAGNATTTCTTAAGATAGCANCCCCCTTNCCTCA 2258  
Db 435 TGTGACTCTCAAACCTCCT---CCAGATTTCTTAATAGCAACCA-----TTCCCTCA 485  
  
Qy 2259 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTTANCTCAGGACTCAGCTAGCATTAAT 2318  
Db 486 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTA-CTCAGGACTCATGTAGCATTAAGT 544  
  
Qy 2319 CAGCTGTGNAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTGGGGCAGGGGAGCGCG 2378  
Db 545 CAGCTGTGAAATTGTGAGGGGGTGTACGTAGCCTCAACCTCTGGGGCAGGGGAGCGCG 604  
  
Qy 2379 AGACTCGTGGGAGAGCTCATTTCCACATCTTCCAGAGCAGCTTNGTTCAGCTGTC 2438  
Db 605 AGACTCTTGGGAGAGCTCATTTCCACGCTTTGCCAAGACAGCTTTT-GTCCAGCTGTC 663  
  
Qy 2439 CACATTGAGTCAGACTCTCCCGGGAGAGCGCCGGCCCCCAGACATATAAGAACTGC 2498  
Db 664 CACAGGAGCAGACTGCTCCCGGGAGAGCAGCCCCGGCCCCCAGACATATAAGAACTGC 723  
  
Qy 2499 AGCCTTGGTATCGAGAGTCTGGGTGTAGAGAACTCTT 2537  
Db 724 AGCCTTGGTATGCG--AGTGTGGGTGTAGAGAACTCTT 760

RESULT 15  
US-60-680-473-10785/c

; Sequence 10785, Application US/60680473  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Darbel, Maher  
; APPLICANT: Bandaru, Raj  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary t  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; CURRENT FILING DATE: 2005-05-13  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 10785  
; LENGTH: 761  
; TYPE: DNA  
; ORGANISM: Macaca Fascicularis  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(761)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-680-473-10785

Query Match 9.8%; Score 257.4; DB 15; Length 761;  
Best Local Similarity 59.8%; Pred. No. 1.4e-51;  
Matches 490; Conservative 1; Mismatches 269; Indels 59; Gaps 11;  
  
Qy 1719 GACATGTGCCCTTTGCTGCTGTACCACTCTTCCCGAGCAGCAGCGCCGAGCCCC 1778  
Db 761 GACACATGNNNNNGCTGCTGGCTATCANNNTCCCGAGCGCAGCAGNNNNNGAGCCCC 702  
  
Qy 1779 TTCAGGCCAGCACTGCCCGAGACTCGCTGGCACTCAGTTCCTCATCTCTAAAGGTGAA 1838  
Db 701 TTCAGGCCAG--NN 644  
  
Qy 1839 GGGTGATGCAGGATATCCCTGCAGGAACAGTCTGTGGATGGACATGATCAGTGTCTNAAG 1898  
Db 643 NNN 586  
  
Qy 1899 GNAAGCAGCAGAGAGAGCGYTCCGGCGGCCCGAGNCCCACTNATCAGTGTNCCAGCGT 1958  
Db 585 GGAAGCAGCAGCAGAGAGCGTCCGGCACCCAGCCCC-----ACTGTGAGGTCCAG 532  
  
Qy 1959 GCTNGGTTNCCCGAGNAGCAGCTNAGNATCANCACTGACACTNCACTTNGCCCTG 2018  
Db 531 CGTGGCGGCCCGCAGAGCAGCTCACGCA-----ACACTCACCTG 488  
  
Qy 2019 CCCTNGGCCANGAGGTACTGCCGACGCACTTTGCACNTCTGATGNACCTCAAAGCA 2078  
Db 487 CCGCGCCCGCGGCGGAGGGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 428  
  
Qy 2079 CTTTCATGGCTNGCCCTCTNNGCAGGNCAGGNCAGTGCAGTGTAGNAG 2138  
Db 427 CTTTCATGGCTG---CCCTCTGGCAGGGCAGGTGAGTGTCTGTAGGAGC----- 380  
  
Qy 2139 CATANGCAAGCCAGGAGATGGGGTGNAAAGGGANCAAGTCTTTGAGCTGTCCANCAATGCA 2198  
Db 379 -----ANNNGCA 328  
  
Qy 2199 TGTGACTNCTCAAACCTCTTNNCCAGNATTTCTTAAGATAGCANCCCCCTTNCCTCA 2258  
Db 327 TGTGACTCTCAAACCTCCT---CCAGATTTCTTAATAGCAACCA-----TTCCCTCA 277  
  
Qy 2259 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTTANCTCAGGACTCAGCTAGCATTAAT 2318  
Db 276 TTGCCCCAGCTTAGCCTCTTCTCCAGGGAGCTA-CTCAGGACTCATGTAGCATTAAGT 218  
  
Qy 2319 CAGCTGTGNAATCGTTCAGGGGTGTCTGCTAGCCTCAAACCTCTGGGGCAGGGGAGCGCG 2378  
Db 2319 CAGCTGTGNAATCGTTCAGGGGTGTCTGCTAGCCTCAAACCTCTGGGGCAGGGGAGCGCG 2378

Mon Jun 13 13:27:13 2005

Db 217 CAGCTGTGAATTGTCAAGGGGGTGTCAAGCTAGCCTCAACCTCCTTGGGGCGGGGACGCCG 158  
QY 2379 AGACTCCGTGGAGAGCTCAATCCACATCTTGGCCAAGACAGCCCTTTNGTCCAGCTGTC 2438  
Db 157 AGACTCCTTGGGAGAGCTCATCCCAAGCTTGGCCAAGACAGCCCTTT-GTCCAGCTGTC 99  
QY 2439 CACATTGAGTCAGACTGTCTCCCGGGGAGAGAGCCCGCCCGCCAGCACATAAAGAACTGC 2498  
Db 98 CACAGCGAGCCAGACTGTCTCCCGGGGAGAGAGCCCGCCCGCCAGCACATAAAGAACTGC 39  
QY 2499 AGCCTTGGTACTGACAGCTCTGGGTTGTAGAGAACTCTT 2537  
Db 38 AGCGTTGGTACTGC--AGTGTGGGTTGTAGAGAACTCTT 2

Search completed: June 11, 2005, 05:07:11  
Job time : 1021.45 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:18:24 ; Search time 3562.59 Seconds  
(without alignments)  
17042.190 Million cell updates/sec

Title: US-09-155-676B-4  
Perfect score: 1253  
Sequence: 1 cattggagtcacgcggtggc.....ggggggcggtaccacatttt 1253

Scoring table: IDENTITY NUC  
Gapop 10'0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.on.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.by.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	99.9	1253	6	A66648
2	1021.8	81.5	1202	9	AB094093 Sequence 4
3	723.8	57.8	1175	10	BC055929 Mus muscu
4	489	39.0	104913	9	AL109583 Human DNA
5	489	39.0	192748	2	AC011089
6	444.4	35.5	500	6	CQ700802
7	400	31.9	601	11	BV167722
8	400	31.9	1191	11	BV177918
9	392	31.3	401	6	AX071560
10	327.4	26.1	486	6	CQ752826
11	309	24.7	1390	5	CR406372
12	297	23.7	421	6	BD058555
13	281.4	22.5	433	6	CQ727443
14	263	21.0	263	6	CQ728296
15	244.8	19.5	374	6	AX906736
16	244.8	19.5	374	6	BD042269
17	186.4	14.9	254	6	CQ707871
18	175.8	14.0	158634	2	AP000909
19	175.8	14.0	165394	2	AP000877 Homo sapi

20	175.8	14.0	178341	9	AP003392
21	175.8	14.0	196511	2	AP000854
c 22	175.8	14.0	244847	2	AP002357
c 23	173	13.8	165394	2	AP000877
c 24	173	13.8	196511	2	AP000854
c 25	173	13.8	244847	2	AP002357
c 26	160.8	12.8	3793	9	AB061844
c 27	134.6	10.7	246370	2	AC107575
c 28	134.6	10.7	248500	2	AC104400
c 29	134.6	10.7	258426	2	AC105645
30	132.6	10.6	185902	10	AC122428
31	122.8	9.8	328	6	AX918334
32	122.8	9.8	328	6	BD053867
33	103.4	8.3	1727	5	BX933109
34	102.4	8.2	158634	2	AP000909
35	83.4	6.7	97	6	AX899765
36	83.4	6.7	97	6	BD035298
c 37	67.6	5.4	701	10	BC079541
c 38	61.6	4.9	1906	6	A66645
c 39	60.4	4.8	7218	6	I66494
c 40	55.2	4.4	125020	9	AF429315
c 41	54.6	4.4	2000	6	AX655393
42	53.2	4.2	2631	6	A66647
c 43	50.8	4.1	198378	2	AC108256
c 44	49	3.9	97839	2	AC141520
45	47.8	3.8	1080	9	AF035606

ALIGNMENTS

RESULT\_1  
LOCUS A66648 1253 bp DNA linear PAT 29-MAR-1999  
DEFINITION Sequence 4 from Patent WO9737016.  
ACCESSION A66648  
VERSION A66648.1 GI:4538140  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1253)  
AUTHORS Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.  
TITLE MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR PREPARATION AND USE  
JOURNAL Patent: WO 9737016-A 4 09-OCT-1997;  
COMMENT YEDA RES & DEV (IL)  
FEATURES Other publication AU 2175597 19971022.  
source Location/Qualifiers  
1..1253  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

ORIGIN

Query Match 99.9%; Score 1252; DB 6; Length 1253;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CATTGGAGTCAACGCGGTGGCGGCGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60  
Db 1 CATTGGAGTCAACGCGGTGGCGGCGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60  
Qy 61 GATTTCGAGCCCAAGAGGCCCTTCTTCTGTGGTCGGGCACGTTTACAGCCGCAAGCAC 120  
Db 61 GATTTCGAGCCCAAGAGGCCCTTCTTCTGTGGTCGGGCACGTTTACAGCCGCAAGCAC 120  
Qy 121 CCAGCGCAGCTGAAGGAGGCTTTTGAGAGGCTCTCTCCCGAGGTGAGGCGGCCCGCAA 180  
Db 121 CCAGCGCAGCTGAAGGAGGCTTTTGAGAGGCTCTCTCCCGAGGTGAGGCGGCCCGCAA 180  
Qy 181 GGCCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGAAACACGAGCGATCTCTGTGTG 240  
Db 181 GGCCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCCGAAACACGAGCGATCTCTGTGTG 240

Db 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGAAACACGAGCGATGCTGCTGGT 240  
Qy 241 CCTGTGCTGCGGCTGTGAGTGCAGGAAACCTGAGCCATGCAACCTGAGCGTCTGTA 300  
Db 241 CCTGTGCTGCGGCTGTGAGTGCAGGAAACCTGAGCCATGCAACCTGAGCGTCTGTA 300  
Qy 301 CGGGGGGCTGTGGAGCTCTGGCCAGCCAGAGCACCAAGAACCAACCAAAATTCG 360  
Db 301 CGGGGGGCTGTGGAGCTCTGGCCAGCCAGAGCACCAAGAACCAACCAAAATTCG 360  
Qy 361 GTGGAGAACAAAGCTGAGTCCAGATGAAGAGAGTTCCTGCTCATCTCCAGGATTA 420  
Db 361 GTGGAGAACAAAGCTGAGTCCAGATGAAGAGAGTTCCTGCTCATCTCCAGGATTA 420  
Qy 421 TCGCGATTCAAGAAATCATGTCGAAAGTTTGGATTCCTATGAAGAAAGGAGATAA 480  
Db 421 TCGCGATTCAAGAAATCATGTCGAAAGTTTGGATTCCTATGAAGAAAGGAGATAA 480  
Qy 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGTGGT 540  
Db 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGTGGT 540  
Qy 541 TCGGCTGTCTTAGAGCTCAGGCAGTGCAGACCCAGAGAGGCTCTTCAGACCTTAG 600  
Db 541 TCGGCTGTCTTAGAGCTCAGGCAGTGCAGACCCAGAGAGGCTCTTCAGACCTTAG 600  
Qy 601 AGCTGGAAAGGATGAACAGCAAGTACGCTTCCAGCTTACAGAGCCCTCAAAATTTGA 660  
Db 601 AGCTGGAAAGGATGAACAGCAAGTACGCTTCCAGCTTACAGAGCCCTCAAAATTTGA 660  
Qy 661 CTGCCCACAGCTCCAGAGCTTGAATGATGGAGAGCAGGACCATCTCTGACATTCATTGG 720  
Db 661 CTGCCCACAGCTCCAGAGCTTGAATGATGGAGAGCAGGACCATCTCTGACATTCATTGG 720  
Qy 721 CCATCAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACTCCCTCGGATGAT 780  
Db 721 CCATCAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACTCCCTCGGATGAT 780  
Qy 781 CCAAGTGAAGAAATCATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTC 840  
Db 781 CCAAGTGAAGAAATCATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTC 840  
Qy 841 TAAAGAAAAGGAAAACAGAAAGTTGAAAACCTCCCGCCAGACCGAGTTGGGCGCAACTT 900  
Db 841 TAAAGAAAAGGAAAACAGAAAGTTGAAAACCTCCCGCCAGACCGAGTTGGGCGCAACTT 900  
Qy 901 TGATCAGCTCCAGGACGAGTGCAGGCTGGCTGCGCTCTTTTGGCGCGCTCTGGAATA 960  
Db 901 TGATCAGCTCCAGGACGAGTGCAGGCTGGCTGCGCTCTTTTGGCGCGCTCTGGAATA 960  
Qy 961 ATGGAGCGCGCTGGAGTCCAGACATCACTCCAAACTGAAGCTGCAGCAATGAAGAG 1020  
Db 961 ATGGAGCGCGCTGGAGTCCAGACATCACTCCAAACTGAAGCTGCAGCAATGAAGAG 1020  
Qy 1021 CAGTCAATACAGAAAAAGCTTAATCATGCTCTCTACCAACTACCATGAGGCTAAAGCC 1080  
Db 1021 CAGTCAATACAGAAAAAGCTTAATCATGCTCTCTACCAACTACCATGAGGCTAAAGCC 1080  
Qy 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCCAAATTCCTTATCATCTTCTTAG 1140  
Db 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCCAAATTCCTTATCATCTTCTTAG 1140  
Qy 1141 GAAACAGATACCTCATTTGATTTAATAAGTTTATTTTTCGGCTTCGCGCT 1200  
Db 1141 GAAACAGATACCTCATTTGATTTAATAAGTTTATTTTTCGGCTTCGCGCT 1200  
Qy 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGGGGGGCGCTACCCACTTTT 1253  
Db 1201 CGAATCAAGCTTATCGATACCGTCCAGCTCGAGGGGGGCGCTACCCACTTTT 1253

LOCUS AB094093 1202 bp mRNA linear PRI 02-MAY-2003  
DEFINITION Homo sapiens DLNB14 mRNA, complete cds.  
ACCESSION AB094093  
VERSION AB094093.1 GI:30314480  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 Kubo,T., Arai,Y., Ohira,M., Gamou,T., Maeno,G., Sakiyama,T., Toyoda,A., Hattori,M., Sakaki,Y., Nakagawara,A. and Ohki,M. Identification of a 500-kb region of common allelic loss in chromosome 11q23 in non-MYC amplified type of neuroblastoma Unpublished  
2 (bases 1 to 1202)  
Arai,Y., Kubo,T., Ohira,M., Gamou,T., Maeno,G., Sakiyama,T., Toyoda,A., Hattori,M., Nakagawara,A. and Ohki,M. Submitted (17-OCT-2002) Yasuhiro Arai, National Cancer Center Research Institute, Molecular Oncology Division; 5-1-1 Teukiji, Chuo-ku, Tokyo 104-0045, Japan (E-mail: yaraiegan2.ncc.go.jp, Tel:81-3-3542-2511(ex.4752), Fax:81-3-3542-0688)  
This gene was identified within a LOH region at 11q23.3 of neuroblastoma. Reference sequence was predicted from genomic sequences of a BAC contig mapped at 11q23.3 using several gene prediction programs. Expression was analyzed by RT-PCR and Northern blot. Then, a part of this cDNA was obtained by cDNA library screening.  
FEATURES  
Location/Qualifiers  
1..1202  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11q23.3"  
1..1202  
/gene="DLNB14"  
57..1055  
/gene="DLNB14"  
/codon\_start=1  
/protein\_id="BAC76047.1"  
/translation="GI:30314481"  
/db\_xref="GI:30314481"  
RKAARAAQVRYVPEHEKRCWCLCGCEVRHLSHNLTVLYGLLEHLAGPEHKKAT NKFWENKAEVQMEKETLVPQDYARFKKSMVKGLDSVEEKDKVIEKMAQIRVEQ SRQEVRSVLFPQAVPPEEGSSAPRSWKMNQSVASSLOQPSNLDLPAPELDMET GPSLTFIGHQDIPGVNIHSGATFPWMIQDEYTAGNOEIGSVSEEFLEKEKEKQKLKK LPDDRVGANFHDSSRTSAGWLPFSGRVNNGRRWQSRHOFKTEAAAMKKQSHTKS"  
ORIGIN  
Query Match 81.5%; Score 1021.8; DB 9; Length 1202;  
Best Local Similarity 98.7%; Pred. No. 5.5e-260;  
Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;  
Qy 81 CTTCTCTTCTGTGTCGC-GGCACGTTTACAGCCGCAAGCACCCAGCGCAGCTGAAGGAG 139  
Db 94 CTTCTCTTCTGTGTCGCGGCACGTTTACAGCCGCAAGCA-CCAGCGCAGCTGAAGGAG 152  
Qy 140 GCTTTTGAAGGCTCTGCCCCAGGTGAGGCGGCCCGCAAGGCCATCGCGCCGCTCAG 199  
Db 153 GCTTTGAAGGCTCTGCCCCAGGTGAGGCGGCCCGCAAGGCCATCGCGCCGCTCAG 212  
Qy 200 GTGGAGCGCTATGTCGCCGAAACACGAGCGATGCTCTGTCGCTGCTGTCGGCTGTGAG 259  
Db 213 GTGGAGCGCTATGTCGCCGAAACACGAGCGATGCTCTGTCGCTGCTGTCGGCTGTGAG 272  
Qy 260 GTGCGGGAACACCTGAGCCATGGAACCTGACGCTCTGTACGGGGGCTGCTGGAGCAT 319  
Db 273 GTGCGGGAACACCTGAGCCATGGAACCTGACGCTGCTGTACGGGGGCTGCTGGAGCAT 332  
Qy 320 CTGGCCAGCCGACGACCAAGAAACCAACAAATTCGTGGGAGNACAAAGCTGAG 379

Db	333	CTGGCCAGCCGACGACACAGAAACCAACCAAAATTTCTGTTGGGAGAACAAAGCTGAG	392
Qy	380	GTCCAGATGAAGAAGATTTCTGTCTACTCCCCAGGATTATGCGCGATTCAAGAAATCC	439
Db	393	GTCCAGATGAAGAAGATTTCTGTCTACTCCCCAGGATTATGCGCGATTCAAGAAATCC	452
Qy	440	ATGTGAAAGGTTTGATTTCTATGAAGAAGAGGATTAAGTGATCAAGGAGATGGCA	499
Db	453	ATGTGAAAGGTTTGATTTCTATGAAGAAGAGGATTAAGTGATCAAGGAGATGGCA	512
Qy	500	GCTCAGATCCGTGAGGTCGAGCAGAGCCGACAGAGAGTGGTTTCGGTCTGTCTTAGAGCCT	559
Db	513	GCTCAGATCCGTGAGGTCGAGCAGAGCCGACAGAGAGTGGTTTCGGTCTGTCTTAGAGCCT	572
Qy	560	CAGCAGTGCACAGACCCAGAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAAC	619
Db	573	CAGCAGTGCACAGACCCAGAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAAC	632
Qy	620	AGCCAAAGTAGCTTCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACACAGCTCCAGAG	679
Db	633	AGCCAAAGTAGCTTCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACACAGCTCCAGAG	692
Qy	680	CTTGACTCGATGGAGACAGGACCATCTCTGACATTTATGGCCATCAGGATATACCAGGA	739
Db	693	CTTGACTCGATGGAGACAGGACCATCTCTGACATTTATGGCCATCAGGATATACCAGGA	752
Qy	740	GTTGGTAATCACTCACTAGAGTGCCACACCTCCCTGGATGATCCAAAGTGAAGAATACATT	799
Db	753	GTTGGTAATCACTCACTAGAGTGCCACACCTCCCTGGATGATCCAAAGTGAAGAATACATT	812
Qy	800	GCTGGAAACCAAGAAATAGGACATCTCTGACATTTATGAAGATTTCTTAAAGAAAGGAAAAACAG	859
Db	813	GCTGGAAACCAAGAAATAGGACATCTCTGACATTTATGAAGATTTCTTAAAGAAAGGAAAAACAG	872
Qy	860	AAGTTGAAAAAATCCCTCCACAGCGAGTGGGCGCACTTTGATCAGACTCCAGGACC	919
Db	873	AAGTTGAAAAAATCCCTCCACAGCGAGTGGGCGCACTTTGATCAGACTCCAGGACC	932
Qy	920	AGTGACAGCTGGCTGCTCTTTTGGGCGCGGCTCTGGAATATGAGACCGCCCTGGCAGTC	979
Db	933	AGTGACAGCTGGCTGCTCTTTTGGGCGCGGCTCTGGAATATGAGACCGCCCTGGCAGTC	991
Qy	980	CAGACATCACTCCAAAATGAGCTGAGCAATGAAGAAGAGTGCATATCAGAAAAA	1039
Db	992	CAGACATCAA-TTCAAAAATGAGCTGAGCAATGAAGAAGAGTGCATATCAGAAAAA	1050
Qy	1040	GCTAATCATGCTCTTACCACTACCATGAGCTAAAGCCAAAGTCAACCAACCCCTA	1099
Db	1051	GCTAATCATGCTCTTACCACTACCATGAGCTAAAGCCAAAGTCAACCAACCCCTA	1108
Qy	1100	TTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTTAGGAAACAGACATATCTCATTC	1159
Db	1109	TTATACCTTCCA-CCAAATTTCTTATCATTTCTTTCTTAGGAAACAGACATATCTCATTC	1167
Qy	1160	ATTGATTTAATAAAGTTTATTTTC	1186
Db	1168	ATTGATTTAATAAAGTTTATTTTC	1194
RESULT 3			
BC055929			
LOCUS			
DEFINITION	Mus musculus RIKEN cDNA D63004F24 gene, mRNA	linear	ROD 25-MAR-2004
ACCESSION	BC055929		
VERSION	IMAGE:4205568), partial cds.		
KEYWORDS	BC055929.1	GI:33416549	
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 1175)		
	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,		

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
12477932	
2 (bases 1 to 1175)	
Strausberg, R.	
Direct Submission	
Submitted (01-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>	
Contact: MGC help desk	
Email: <a href="mailto:cgapsb-remail.nih.gov">cgapsb-remail.nih.gov</a>	
Tissue Procurement: Jeffrey E. Green, M.D.	
cDNA Library Preparation: Life Technologies, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305	
Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a>	
Contact: (Dickson, Mark) <a href="mailto:mcd@paxil.stanford.edu">mcd@paxil.stanford.edu</a>	
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>	
Series: IRAK Plate: 120 Row: j Column: 5	
This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.	
FEATURES	
source	
1..1175	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="FVB/N"	
/db_xref="taxon:10090"	
/clone="IMAGE:4205568"	
/tissue_type="Colon, normal. 5 month old male mouse."	
/clone_lib="NCI CGAP_C024"	
/lab_host="DH10B"	
/note="vector: pCMV-SPORT6"	
<1..1175	
/gene="D63004F24R1k"	
/db_xref="LocusID:382073"	
/db_xref="MGI:2685960"	
<1..1018	
/gene="D63004F24R1k"	
/codon_start=2	
/product="D63004F24R1k protein"	
/protein_id="AAH55929.1"	
/db_xref="GI:33416550"	
/db_xref="LocusID:382073"	
/db_xref="MGI:2685960"	
/translation="FPGFSMAPPPQRCPLCRQTFFCGRHVYSHHQRLKGALERLLL PQYEAARRVAAQVRYVEPHDRCCWCPCCGCEVRKHLSHGNLTVLHGGLEHLASP EHKKATNKFVWENKANAKMEKEFLISPDYARFKSKVMKGLDSYEEKEDEVIKMAAQ	

IREVBSQREVVRSLVLEPQABSDPEGSSAPESKWKATNGHVASSSQVSHLALQPVAE  
LDWETGQOLFLGHODTPEGIGNHSGATPWWLQERHSSGSLPIGPSYEFPLKEXE  
KOKLKKLPDRVGANFDHSSNTSAGWLPFSFGRVWNNRWSRQHFKEATRSKQPR  
KGSK

ORIGIN		Query Match	Score	723.8	DB 10	Length	1175
		Best Local Similarity	81.4%	Pred. No. 7.9e-181			
		Matches	901	Conservative	0	Mismatches	197
						Indels	9
						Gaps	5
QY	81	CTTTCTTCTGTGTGTCGCGG-CACGTTTACAGCGCAAGCACCCAGCGCAGCTGAAGGAG	139				
Db	57	CTTTCTTCTGTGTGTCGCGGTCACGTCACAGGCAACAGCA-CCAGCGCAGTTGAAGGGG	115				
QY	140	GCTTTTGAGAGGCTCTGCCAGGTGGAGGGGGCCGCAAGGCCATCGCGCGCTCAG	199				
Db	116	GCTCTGGAGCGCTTCTACCGCAGGTGGAGGGCGCGCGCGCTCGCGCGCCCGCAG	175				
QY	200	GTGGAGCGCTATGTGCCCCGAACACAGCGCATGCTGTGTGCTGTGTGTGCGGCTGTGAG	259				
Db	176	GTGGAGCGCTATGTGCCCCGAGCAGATCGTGTGTGTGTGTGCGGCTGTGCGAA	235				
QY	260	GTGCGGAAACACTGAGCCATCGAAACCTGACGGTGTGTGACGGGGGCTGTGAGGAT	319				
Db	236	GTGCGAAACACTGAGCCATCGCAACCTGACAGTCTACACGGGGTCTGTGTGAACAT	295				
QY	320	CTGGCCAGCCGAGCAGACAGAAAGCAACCAAAATCTGTGTGGAGAACAAAGCTGAG	379				
Db	296	CTGGCCAGCCGAGCAGACAGAAAGCCCAACAAATCTGTGTGGAGAACAAAGCCCAT	355				
QY	380	GTCCAGATGAAGAGAGTTTCTGTCTCACTCCCGAGGATTATCGCGATTTCAGAAATCC	439				
Db	356	GCCAGATGAAGAGAGTTTCTGTCTCACTCCCGAGGATTATCGCGATTTCAGAAATCC	415				
QY	440	ATGGTGAAGGTTTGATTCCTATGAAGAAAGAGAGATAAAGTGAATCAAGAGATGCA	499				
Db	416	ATGGTGAAGGTTTGATTCCTATGAAGAAAGAGAGATGAATGAATGAAGAGATGCA	475				
QY	500	GCTCAGATCGTGAGTGAGCAGAGCCGACAGGAGGTGTGTGTGTCTCTTAGAGCT	559				
Db	476	GCTCAGATCGTGAGTGAGCAGAGTCGCGAGAGTGTGTGTGTCTCTTAGAGCT	535				
QY	560	CAGCGAGTCCAGACCCAGAGAGGCTCTTTCAGCACCTTAGAGCTGGAAGGATGAAC	619				
Db	536	CAGCGAGTCCAGATCCAGAGAGGCTCTTTCAGCACCTTAGAGCTGGAAGGATGAAC	595				
QY	620	AGCAAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGACCTGCACAGCTCCAGAG	679				
Db	596	GGCCATGTAGCTTCCAGCTTACAGAGGATATCACTTTGGCTTGCACCTGTTCAGAG	655				
QY	680	CTTGACTGGATGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGA	739				
Db	656	CTTGATTTGGATGAGACAGGACGACACTAACATTCATTGGCCATCAGGATATACCGGA	715				
QY	740	GTTGGTAACATCCACTCAGTGGCCACACTCTCTGTGATGATCCAGATGAAGATACAT	799				
Db	716	ATTGGTAACATTCACTCAGTGGCCACACTCTCTGTGATGATCCAGAGGAGGACAGC	775				
QY	800	GCTGGGNACCAAGAAATAGACACCATCTATGAGAAATTTCTTAAGAAAGGAAACAG	859				
Db	776	TCTGGCAGCTCCCAATAGACACCTCTATGAGAAATTTCTTAAGAAAGGAAACAA	835				
QY	860	AAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACC	919				
Db	836	AACTGGAAGAACTCCCTCCAGATAGATTGGCGCCAACTTTGACACAGTTCACACACC	895				
QY	920	AGTGCAGGCTGGTGGCCCTCTTTTGGCGCGGCTCTGGAATTAATGGAGCGCGCTGGCAGTC	979				
Db	896	AGTGCAGGCTGGTGGCCCTCTTTT-GGCAGAGTCTGGAATAATGGAGCGCGCTGGCAGTC	954				
QY	980	CAGACATCAACTCCAAACTGAGGCTCGACGATGAGAGGAGTGCATACATACAGAAAAA	1039				
Db	955	CAGGCATCAA-TTCAAAACCGAAGCTGCAACAGAGCAAGCAACCACTGAAGGGAATA	1013				

QY	1040	GCTAATCATGCTCTCTACCACTACCATGAGGTAAAGCCAAAGTCAACCAACCCCTA	1099
Db	1014	GCTCAAGGTTCCAGGCGCAATTCATCTGACCTCATGTGCGCAACAAACCAATTCAGCTA	1073
QY	1100	TTATACCTTCCACCCAAATCTTTATCATCTCTTTCTTTAGGAAACAGACATCTCTTC	1159
Db	1074	TTATCTCTAAGACCAAA-----TCCTCCCGGCTGCTTAGGAAACAGATGCACCCATC	1128
QY	1160	ATTGATTAAATAAGTTTATTTTTC	1186
Db	1129	ATTGATTCAATAAGTTTATTTTTC	1155

RESULT 4

HSJ718P11/c	104913 bp	DNA	linear	PRI 15-MAR-2001
LOCUS	Human DNA sequence from clone RP4-718P11 on chromosome			
DEFINITION	20p12.1-12.3. Contains part of the gene for a novel class II			
	aminotransferase similar to serine palmitoyltransferase, ESTs, STSS			
	and GSSs, complete sequence.			
ACCESSION	AL109983			
VERSION	AL109983.2 GI:9368492			
KEYWORDS	HTG; aminotransferase; serine palmitoyltransferase.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 104913)			
AUTHORS	Matthews,L.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,			
	CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk			
	requests: clonerequest@sanger.ac.uk			

On Jul 22, 2000 this sequence version replaced gi:5777578.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information  
on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/c\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/c_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 20, constructed by the Sanger Centre Chromosome 20  
Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP4-718P11 The true  
left end of clone RP5-107712 is at 71329 in this sequence. The true  
right end of clone RP11-303K20 is at 39954 in this sequence. This  
sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. RP4-718P11 is from  
the library RPCI-4 constructed by the group of Pieter de Jong. For  
further details see  
<http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

FEATURES	Location/Qualifiers
Source	1..104913
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosomes="20"
	/map="p12.1-12.3"
	/clone="RP4-718P11"
	/clone_lib="RPCI-4"

```

repeat_region 10. 329
repeat_region /note="MER87 repeat: matches 211. .540 of consensus"
repeat_region 544. .611
repeat_region /note="MERSA repeat: matches 121. .189 of consensus"
repeat_region 824. .852
repeat_region /note="MERSB repeat: matches 150. .178 of consensus"
repeat_region 865. .1291
repeat_region /note="L2 repeat: matches 307. .860 of consensus"
repeat_region 1565. .1604
repeat_region /note="10 copies 4 mer ttat 100% conserved"
repeat_region 1643. .2345
repeat_region /note="L1PB3 repeat: matches 5450. .6150 of consensus"
repeat_region 2534. .2687
repeat_region /note="L2 repeat: matches 833. .1002 of consensus"
repeat_region 2704. .3333
repeat_region /note="L2 repeat: matches 1774. .2440 of consensus"
repeat_region 3898. .4038
repeat_region /note="L1MB7 repeat: matches 5776. .5920 of consensus"
repeat_region 4108. .4463
repeat_region /note="L1MB7 repeat: matches 5303. .5661 of consensus"
repeat_region 4829. .5139
repeat_region /note="AluSx repeat: matches 4. .312 of consensus"
repeat_region 5208. .5388
repeat_region /note="AluSg/x repeat: matches 127. .309 of consensus"
repeat_region 5435. .5486
repeat_region /note="26 copies 2 mer tt 84% conserved"
repeat_region 5489. .5786
repeat_region /note="AluSx repeat: matches 1. .309 of consensus"
repeat_region 6844. .6959
repeat_region /note="58 copies 2 mer tt 62% conserved"
repeat_region 6978. .7364
repeat_region /note="L1MA9 repeat: matches 5894. .6275 of consensus"
repeat_region 7354. .7682
repeat_region /note="L1 repeat: matches 4721. .5050 of consensus"
repeat_region 7917. .8004
repeat_region /note="22 copies 4 mer ttcc 63% conserved"
repeat_region 8227. .8292
repeat_region /note="33 copies 2 mer ta 95% conserved"
repeat_region 8229. .8292
repeat_region /note="16 copies 4 mer tata 96% conserved"
repeat_region 10536. .10888
repeat_region /note="MER1B repeat: matches 1. .337 of consensus"
repeat_region 11927. .12045
repeat_region /note="L2 repeat: matches 2143. .2290 of consensus"
repeat_region 12303. .12811
repeat_region /note="MER75 repeat: matches 6. .514 of consensus"
gene 13443. .101318
gene /gene="dJ718P11.1"
CDS
join(<13443. .13597,15536. .15684,20936. .>21054)
/note="match: ESTs: Em:AA192501"
match: ESTs: Em:AA357446 Em:AA192501
Tr:O20375 Tr:Q9XGB0 Sw:Q09925 Sw:P40970 Sw:P48241
/codon_start=1
evidence=not experimental
/product="dJ718P11.1 (novel class II aminotransferase
similar to serine palmitoyltransferase (isoform 1))"
/protein_id="CAB65865.1"
/db_xref="GI:6691980"
/db_xref="GOA:Q9UGB6"
/db_xref="UniProt/TREMBL:Q9UGB6"
/translation="DFVPLVQDFNFYRLNLYMRIRDNWRNPICSAFGPLFLMERVS
DDNWTFRFGVRIKDVINGSNFGLAAYKIDSMRTIKDVLVYGTGVASTRHEMG
TLDKHKLELDVLAFLVAAMVFGMGFATNMMNIPALVGKGCILISDELNHTSLVIG
ARLSGATIRFKNNWTSLEKLDVAVYQPRTRAKKILILVEGVSMESGIVHL
PQIATLKKKAYILYIDRAHSIVAGTGRGVTIEFGLDPEHVDVLMGTFTKFGQAA
GYIAGKLDVLDVSHSVAIVASSMPPIAEQIIRSLKILMDLGTQGLQFGQAA
KNTRYFQRLOEMGFIYIGNENASVWPLLLMPGKVAFAFARMLEKKIIGVVVGFAT
PLAARARFCVSAHTREMLDT"

```

```

CDS
join(<13443. .13597,15536. .15684,20936. .>21054)
/note="dJ718P11.1"
/note="match: ESTs: Em:AA192501"
/codon_start=1
evidence=not experimental
/product="dJ718P11.1.2 (novel class II aminotransferase
similar to serine palmitoyltransferase (isoform 2))"
/protein_id="CAB65866.1"
/db_xref="GI:6691981"
/db_xref="GOA:Q9UGB5"
/db_xref="UniProt/TREMBL:Q9UGB5"
/translation="DFVPLVQDFNFYRLNLYMRIRDNWRNPICSAFGPLFLMERVS
DDNWTFRFGVRIKDVINGSNFGLAAYKIDSMRTIKDVLVYGTGVASTRHEMG
EFTSSWQNALCVGCVLQADSVQPPMSWQXGRHSSLE"
14113. .14582
/note="L1MA4 repeat: matches 5815. .6292 of consensus"
15902. .16080
/note="L2 repeat: matches 2480. .2658 of consensus"
complement(16079. .16250)
/note="match: STS: Em:HSPF04H5"
18972. .19148
/note="AluJo repeat: matches 121. .294 of consensus"
19894. .20269
/gene="dJ718P11.1"
/note="match: GSS: Em:AQ044246"
20181. .20417
/note="L1MA4 repeat: matches 4366. .4597 of consensus"
complement(20538. .20941)
/note="match: GSS: Em:AQ662868"
21219. .21378
/note="L1MA4 repeat: matches 5160. .5326 of consensus"
21415. .21708
/note="AluSc repeat: matches 1. .290 of consensus"
22230. .22469
/note="L1R23 repeat: matches 206. .437 of consensus"
23617. .23863
/note="MERS8B repeat: matches 79. .337 of consensus"
24767. .24827
/note="MIR repeat: matches 63. .123 of consensus"
25619. .25838
/note="MIR repeat: matches 31. .238 of consensus"
25845. .26034
/note="MIR repeat: matches 66. .258 of consensus"
complement(25953. .26556)
/note="match: GSS: Em:AQ482745"
27967. .28282
/note="AluJo repeat: matches 1. .312 of consensus"
28751. .28859
/note="L2 repeat: matches 2620. .2723 of consensus"
29512. .29809
/note="AluY repeat: matches 1. .309 of consensus"
29810. .30116
/note="AluSx repeat: matches 1. .296 of consensus"
30520. .30690
/note="Charlie5 repeat: matches 2414. .2584 of consensus"
30691. .30996
/note="AluSx repeat: matches 1. .306 of consensus"
30997. .31098
/note="Charlie5 repeat: matches 2313. .2414 of consensus"
31097. .31278
/note="Charlie5 repeat: matches -5. .115 of consensus"
31450. .31581
/note="MIR repeat: matches 105. .249 of consensus"
complement(31474. .32053)
/note="match: GSS: Em:AQ083543"
31753. .32181
/note="L1ME3 repeat: matches 5679. .6131 of consensus"
32059. .32525
/gene="dJ718P11.1"
/note="match: GSS: Em:AQ774405"
Query Match 39.0%; Score 489; DB 9; Length.104913;
Best Local Similarity 82.3%; Pred. No. 3.7e-118;

```

Matches		672; Conservative	0; Mismatches	105; Indels	40; Gaps	8;
QY	406	CACTCCCCAGGATTATGCGGATTCAAGAAATCCATGGTCAAGGTTTGATTCCTATGA				
Db	15503					
QY	466	AGAAAAGGAGGATAAAGTGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAG				
Db	15443					
QY	526	CCGACAGGAGGTGGTTCGGTCTGCTCTAGAGCCT-----CAGGCAGTGC				
Db	15383					
QY	570	CAGACCCAGAAGAGGCTCTTCAGCACCTAGAGCTGGAAAGGATGACACGCAAGTAG				
Db	15323					
QY	630	CTTCCAGCTTACAGCAGCCCTCAAATTTGGACCTGCCACGCTCCAGAGCTTGACTGGA				
Db	15263					
QY	690	TGGAGACAGACATCTCTGCATTCATTGGCCATCAGGATATACAGGAGTTGGTAACA				
Db	15204					
QY	750	TCCACTCAGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTGCTGGGACC				
Db	15144					
QY	810	AAGAAATAGGACCATCTATGAAGATTTCTTAAAGAAAGGAAACACAGAAAGTTGMAAA				
Db	15084					
QY	870	AACTCCCCCAGACGAGTTGGGCCAACTTTGATCAACAGTCCAGGACCAATGCAGGCT				
Db	15024					
QY	930	GGCTGCCCTCTTTTGGGCGCTGTGGAATATGACGCGCTGGCAGTCCAGACATCAA				
Db	14964					
QY	990	CTCCAAATGAAGTGCAGCAATGAAGAAGCAGTCA-----CATACAGAAAAAAG				
Db	14905					
QY	1041	CTAATCATGC-----TCTCTACCACTACCATGAGGCTAAAGC-----CAAGTCAAC				
Db	14846					
QY	1090	CAAAACCCCTATTATACCTTCCACCACAAATCTTTATCATTTGTTCTTAGGAAACAGAC				
Db	14786					
QY	1150	ATACCTCATTCATTGATTTAATAAAGTTTATTTTTTCT 1186				
Db	14727					

RESULT 5  
AC011089  
LOCUS  
DEFINITION  
Homo sapiens chromosome 20 clone RP11-303K20 map 20, WORKING DRAFT  
SEQUENCE, 13 unordered pieces.  
AC011089  
AC011089.5 GI:8901222  
VERSION  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 192748)  
AUTHORS  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Homo sapiens chromosome 20, clone RP11-303K20  
Unpublished  
2 (bases 1 to 192748)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,  
Gallagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamae,J.,  
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 192748)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Levine,R., Larocque,K., Lamazares,R., Landers,T., Lehoczky,J.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrim,J., Meneus,D., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Naylor,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamae,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 3, 2000 this sequence version replaced gi:7717100.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L1378  
Center clone name: 303\_K\_20  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 182218 bases at least Q40  
Consensus quality: 187129 bases at least Q30  
Consensus quality: 189385 bases at least Q20  
Insert size: 194000; agarose-fp  
Insert size: 191548; sum-of-contigs  
Quality coverage: 4.6 in Q20 bases; agarose-fp  
Quality coverage: 4.7 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 13 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as



Chondrogene Inc. (CA)  
Location/Qualifiers  
1. .500  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

FEATURES  
source

ORIGIN  
Query Match 35.5%; Score 444.4; DB 6; Length 500;  
Best Local Similarity 97.8%; Pred. No. 1.1e-106;  
Matches 493; Conservative 0; Mismatches 6; Indels 5; Gaps 4;  
1 GACTGGATGAGACAGAGACCATCTCTGACATTCATTGGCCATCAGGATATACAGAGTT 742  
1 GACTGGATGAGACAGAGACCATCTCTGACATTCATTGGCCATCAGGATATACAGAGTT 60  
743 GGTAAATCATCAGTCCAGTCCACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCT 802  
61 GGTAAATCATCAGTCCAGTCCACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCT 120  
803 GGGAAACCAAGATAGGACCATCTATGAAGATTTCTTAAAGAAAGGAAAAACAGAG 862  
121 GGGAAACCAAGATAGGACCATCTATGAAGATTTCTTAAAGAAAGGAAAAACAGAG 180  
863 TTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAT 922  
181 TTGAAAAAATCTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAT 240  
923 GCAGGTGGTGGCTCTTTTGGCGCGCTCTGGAAATATGAACCGCGCTGCAGTCCAG 982  
241 GCAGGTGGTGGCTCTTTTGGCGCGCTCTGGAAATATGAACCGCGCTGCAGTCCAG 299  
983 ACATCAACTCCAAACTGAAGCTGCAGCAATGAAGCAGTCAATACAGAAAAAGCT 1042  
300 ACATCAA-TTCAAACTGAAGCTGCAGCAATGAAGCAGTCAATACAGAAAAAGCT 358  
1043 AATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAGTCAACAAACCCCTATTA 1102  
359 AATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC-AAAGTCAACAAACCCCTATTA 416  
1103 TACCTTCCACCAATCTTTTATCATGTCCTTTCTTAGGAAACAGACATACATTCATT 1162  
417 TACCTTCCA-CCAAATCTTTTATCATGTCCTTTCTTAGGAAACAGACATACATTCATT 475  
1163 TGATTTAATAAAGTTTATTTTTC 1186  
476 TGATTTAATAAAGTTTATTTTTC 499

RESULT 7  
BV167722  
LOCUS  
DEFINITION  
sqm6310 Human DNA (Sequencem) Homo sapiens STS genomic, sequence  
tagged site.  
BV167722  
VERSION  
BV167722.1 GI:48001329  
KEYWORDS  
STS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 601)  
Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequencem, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018

BV167722 601 bp DNA linear STS 10-JUN-2004  
sqm6310 Human DNA (Sequencem) Homo sapiens STS genomic, sequence  
tagged site.  
BV167722  
VERSION  
BV167722.1 GI:48001329  
KEYWORDS  
STS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 601)  
Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequencem, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018

Fax: 18582029020  
Email: abraun@sequencem.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 601.  
Location/Qualifiers  
1. .601  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone.lib="Human DNA (Sequencem)"  
<1. .>601

FEATURES  
source

ORIGIN  
Query Match 31.9%; Score 400; DB 11; Length 601;  
Best Local Similarity 97.0%; Pred. No. 7.3e-95;  
Matches 448; Conservative 1; Mismatches 8; Indels 5; Gaps 4;  
725 CAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCAA 784  
77 CAGGATATACAGGAGTTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCAA 136  
785 GATGAAGATATACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCTTAAA 844  
137 GATGAAGATATACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCTTAAA 196  
845 GAAAGGAAAAACAGAGTTGAAATAAATCTCCCCCAGACCGAGTTGGGGCCAACTTTGAT 904  
197 GAAAGGAAAAACAGAGTTGAAATAAATCTCCCCCAGACCGAGTTGGGGCCAACTTTGAT 256  
905 CACAGCTCCAGGACCGAGTGGCTGGCTGGCTTTTGGGGCCGCTCTGGAATAATGG 964  
257 CACAGCTCCAGGACCGAGTGGCTGGCTGGCTTTTGGGGCCGCTCTGGAATAATGG 315  
965 ACGCCGTGGCAGTCCAGACATCAACTCCAAATCTGAAGCTGCAGCAATGAAGACAGT 1024  
316 ACGCCGTGGCAGTCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATGAAGACAGT 374  
1025 CACATACAGAAAAAGCTAATCATGCTCTTACCAACTACCATGAGGCTAAAAGCCAAAG 1084  
375 CACATACAGAAAAAGCTAATCATGCTCTTACCAACTACCATGAGGCTAAAAGC--AAA 432  
1085 TCACCAAAACCCCTATTATACCTTCCACCCAAATCTTTTATCATGTCCTTTCTTAGGAAA 1144  
433 GTCAACAAACCCCTATTATACCTTCCA-CCAAATCTTTATCATGTCCTTTCTTAGGAAA 491  
1145 CAGACATACCTCATTCATTGATTTAATAAAGTTTATTTTTC 1186  
492 CAGACATACCTCATTCATTGATTTAATAAAGTTTATTTTTC 533

RESULT 8  
BV177918  
LOCUS  
DEFINITION  
sqm97502 Human DNA (Sequencem) Homo sapiens STS genomic, sequence  
tagged site.  
BV177918  
VERSION  
BV177918.1 GI:48014136  
KEYWORDS  
STS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1191)  
Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
Genome Res. (2004) In press  
Contact: Andreas Braun  
Pharmaceuticals division  
Sequencem, Inc.



3595 John Hopkins Court, San Diego, CA 92121, USA

Tel: 18582029018

Fax: 18582029020

Email: abraun@sequenom.com

Primer A: No primer sequence submitted

Primer B: No primer sequence submitted

STS size: 1191

Location/Qualifiers

1. .1191

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/clone\_lib="Human DNA (Sequenom)"

<1. .>1191

# FEATURES

source

# STS

## ORIGIN

```
Query Match      31.9%; Score 400; DB 11; Length 1191;
Best Local Similarity 97.0%; Pred. No. 8.1e-95;
Matches 448; Conservative 1; Mismatches 8; Indels 5; Gaps 4;

Qy 725 CAGGATATACAGGAGTTGGTAACATCCACTCAGTGCCACACCTCCCTGGATGATCCAA 784
Db 270 CAGGATATACAGGAGTTGGTAACATCCACTCAGTGCCACACCTCCCTGGATGATCCAA 329
Qy 785 GATGAAGTAATACATGCTGGGAACCAAGAAATAGGACCATCCTATGAAGAATTTCTTAAA 844
Db 330 GATGAAGTAATACATGCTGGGAACCAAGAAATAGGACCATCCTATGAAGAATTTCTTAAA 389
Qy 845 GAAAGGAAACAGAGTTCGAAAGAACTCCCGCCAGACCCAGTGGGGCCCACTTTGAT 904
Db 390 GAAAGGAAACAGAGTTCGAAAGAACTCCCGCCAGACCCAGTGGGGCCCACTTTGAT 449
Qy 905 CACAGCTCCAGGACAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTCGAATAATGG 964
Db 450 CACAGCTCCAGGACAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTCGAATAATGG 508
Qy 965 ACGCGCTGGGAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAGCAGT 1024
Db 509 ACGCGCTGGGAGTCCAGACATCAA-TTCAAAAACCTGAAGCTGCAGCAATGAAGAGCAGT 567
Qy 1025 CACATACAGAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCCAAAG 1084
Db 568 CACATACAGAAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCCAAAG 625
Qy 1085 TCAACCAACCCCTATTATACCTCCACCCAAATCTTTATCATGCTCTTTCTTAGGAAA 1144
Db 626 GTCACAAACCCCTATTATACCTTCCA-CCAAATCTTTATCATGCTCTTTCTTAGGAAA 684
Qy 1145 CAGACATACCTCATCTGATTTGATTTAATAAGTTTATTTTTC 1186
Db 685 CAGACATACCTCATTTGATTTAATAAGTTTATTTTTC 726
```

```
RESULT 9
AX071560 LOCUS      401 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 2032 from Patent WO0102568.
ACCESSION AX071560
VERSION AX071560.1 GI:12581911
```

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,

Kassam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,

Lamson, G., Drmanac, R., Crkencak, R., Drmanac, S., Dickson, M.,

Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.

Human genes and gene expression products

Patent: WO 0102568-A 2032 11-JAN-2001;

CHIRON CORPORATION (US); HYSEQ, INC. (US)

Location/Qualifiers

1085 TCAACCAACCCCTATTATACCTCCACCCAAATCTTTATCATGCTCTTTCTTAGGAAA

626 GTCACAAACCCCTATTATACCTTCCA-CCAAATCTTTATCATGCTCTTTCTTAGGAAA

1145 CAGACATACCTCATCTGATTTGATTTAATAAGTTTATTTTTC 1186

685 CAGACATACCTCATTTGATTTAATAAGTTTATTTTTC 726

source

1. .401

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Query Match 31.3%; Score 392; DB 6; Length 401;

Best Local Similarity 98.8%; Pred. No. 9.1e-93;

Matches 395; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 475 GGATAAAGTATCAAGGAGATGGCAGCTCAGATCCGTGAGTGAGCAGACCGACAGGA 534

Db 2 GCACGAGTATCAAGGAGATGGCAGCTCATATCCGTGAGTGAGCAGACCGACAGGA 61

Qy 535 GGTGTTCCGTTCTCTTTAGAGCCTCAGGAGTCCAGACCCAGAGAAGGGCTCTTTCAGC 594

Db 62 GGTGTTCCGTTCTCTTTAGAGCCTCAGGAGTCCAGACCCAGAGAAGGGCTCTTTCAGC 121

Qy 595 ACCTAGAAGCTGGAAGGGATGAACAGCAAGTAGTTCAGCTTACAGAGCCCTCAAA 654

Db 122 ACCTAGAAGCTGGAAGGGATGAACAGCAAGTAGTTCAGCTTACAGAGCCCTCAAA 181

Qy 655 TTTCGACCTCCACAGCTCCAGAGCTTGAAGTGGAGACAGACAGCATTCTTGACATT 714

Db 182 TTTCGACCTCCACAGCTCCAGAGCTTGAAGTGGAGACAGACAGCATTCTTGACATT 241

Qy 715 CATTGGCCATCAGGATATACAGGAGTTGGTAAATCCACTCAGGTGCCACACCTCCCTG 774

Db 242 CATTGGCCATCAGGATATACAGGAGTTGGTAAATCCACTCAGGTGCCACACCTCCCTG 301

Qy 775 GATGATCAAGATGAAGAAATACATTCGTGGGAACCAAGAAATAGGACCATCCTATGAAGA 834

Db 302 GATGATCAAGATGAAGAAATACATTCGTGGGAACCAAGAAATAGGACCATCCTATGAAGA 361

Qy 835 ATTTCCTTAAAGAAAGGAAACAGAAAGTTCGAAAGAACTC 874

Db 362 ATTTCCTTAAAGAAAGGAAACAGAAAGTTCGAAAGAACTC 401

## RESULT 10

LOCUS 486 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 38760 from Patent WO02068579.

ACCESSION CQ752826

VERSION CQ752826.1 GI:42390079

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

TITLE Kits, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses

thereof

JOURNAL Patent: WO 02068579-A 38760 06-SEP-2002;

PE Corporation (NY) (US)

Location/Qualifiers

1. .486

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## FEATURES

source

## ORIGIN

Query Match 26.1%; Score 327.4; DB 6; Length 486;

Best Local Similarity 83.0%; Pred. No. 1.3e-75;

Matches 404; Conservative 0; Mismatches 66; Indels 17; Gaps 2;

Qy 416 GATTATCGCGATTCGAAGAAATCCATGCTGAAGAGTTTGGATTCTCTATGAAGAAAGAG 475

Db 1 GATTATCGCGATTCGAAGAAATCCCTGCTGAAGAGTTTGGTTCTATGAAGAAAGAG 60

Qy 476 GATAAAGTATCAAGGAGATGGCAGCTCAGATCCGTGAGTGAGCAGACCGACAGGAG 535

```
Db      61  GATGAGGTGACCAAGGAGATGGCAGCTCAGATCTGTGAGGTGGAACAGAGCTGGCAGGAG 120
Qy      536 GTGTTTCGGTCTGTCTTAGAGCT-----CAGGCAGTGCACACCAGA 579
Db      121 GTGTTTCAGTCTGTCTTAGAGTGTTCCTCGGAGGATCCAGACCCCTCAGGCAGT 180
Qy      580 AGAGGGCTCTTCAGCACCTAGAGCTGGAAGGATGAACAGCAAGTAGTCTTCCAGCTT 639
Db      181 GCCGGGCACCTTAGTCCCTGGAAGCTGGAAGGATGAACAGCAGGTAGTCTCAGCCC 240
Qy      640 ACAGCAGCCCTCAAAATTGACCTGCCACAGCTCCAGAGCTTGAATGGATGAGACAGG 699
Db      241 ACAGCAGGCTCCAACTTGGACCTGCCA-CAGCTCCAGAGCTTGAATGGATGAGACAGG 299
Qy      700 ACCATCTCTCACATTCAATTCGCCATCAGGATATACAGGAGTGTGTAACATCCACTCAGG 759
Db      300 ACTATCTCCGATATTCAATGCTCATCAGGATATACAGGAGTGTGTAACATCCACTCAGG 359
Qy      760 TGCACACCTCCCTCGATGATCAAGATGAAGATATACATTCGTGGGAACCAAGAAATAGG 819
Db      360 TGCACACCTCCCTCGATGATCCAGGATGAAGATGCATCTCTGGGAACCAACACAGG 419
Qy      820 ACCATCTTGAAGATTTCTTAAGAAAAGGAAAACAGAGTTGAAAAAATCTCCCCC 879
Db      420 ACCGTCTTAAAGATTTATTAAAGAAAAGGAAAAGAGAGTTGAAGAAATTTCCCC 479
Qy      880 AGACCGA 886
Db      480 AACTGA 486
```

```
RESULT 11
CR406372      1390 bp      mRNA      linear      VRT 04-MAY-2004
LOCUS      Gallus gallus finished cDNA, clone CHEST321117.
CR406372
VERSION      CR406372.1      GI:47021906
KEYWORDS
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
              1 (bases 1 to 1390)
              Boardman, P.E., Bonfield, J.K., Brown, W.R.A., Carder, C., Chalk, S.E.,
              Croning, M.D.R., Davies, R.M., Francis, M.D., Grahame, D.V.,
              Hubbard, S.J., Humphray, S.J., Hunt, P.J., Maddison, M., McIaren, S.R.,
              Niblett, D., Overton, I.M., Rogers, J., Scott, C.E., Taylor, R.G.,
              Tickle, C. and Wilson, S.A.
              Direct Submission
              Submitted (04-MAY-2004) Sanger Institute, Hinxton, Cambridgeshire,
              CB10 1SA, UK. E-mail enquiries: chicken@bms.umist.ac.uk
              BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST Gallus gallus cDNA
              sequencing project.
              This sequence is from the
              BBSRC/Dundee/Nottingham/Sanger/Sheffield/UMIST cDNA collection,
              from a library constructed by Elizabeth Bosch. cDNA was prepared
              from RNA extracted from heads, normalised, and poly A-trimmed.
              EcoRI-NotI cut cDNA was then ligated into the vector. Vector:
              pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI Host: Escherichia
              coli DH10B.
FEATURES
            Location/Qualifiers
            1..1390
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /strain="White Leghorn, Hisex"
            /db_xref="taxon:9031"
            /clone="CHEST321117"
            /clone_lib="CSEQCHN24"
            /dev_stage="stage 22"
ORIGIN
```

RESULT 12  
LOCUS BD058555 421 bp DNA linear PAT 27-AUG-2002  
DEFINITION Secreted expressed sequence tags (sESTs).  
ACCESSION BD058555  
VERSION BD058555.1 GI:22604161  
KEYWORDS JP 2001519666-A/410.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 421)  
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,  
Tracy,M., Spaulding,V. and Agostino,M.J.  
TITLE Secreted expressed sequence tags (sESTs)  
JOURNAL Patent: JP 2001519666-A 410 23-OCT-2001;  
GENETICS INSTITUTE INC  
COMMENT PN JP 2001519666-A/410  
PD 23-OCT-2001  
PF 10-APR-1998 JP 1998543068  
PR 10-APR-1997 US 08/835913  
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI  
DAVID MERBERG,  
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC  
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:  
Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
FEATURES  
source Location/Qualifiers  
1..421  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
ORIGIN  
Query Match 23.7%; Score 297; DB 6; Length 421;  
Best Local Similarity 96.8%; Pred. No. 1.6e-67;  
Matches 303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 604 CTGGAAGGATGAACACGCAAGTAGCTTCAGCTTACAGCAGCGCTCAAAATTTGGACCT 663  
Db 100 CTGGCTCATGTTTTCAGCCAAAGTAGCTTCAGCTTACAGCAGCGCTCAAAATTTGGACCT 159  
Qy 664 GCCACGCTCCAGAGCTTGACTGTGATGGACAGACCATCTCTGCACATTCATTGGCCA 723  
Db 160 GCCACGCTCCAGAGCTTGACTGTGATGGACAGACCATCTCTGCACATTCATTGGCCA 219  
Qy 724 TCAGGATATACAGAGTTGGTAAATCACTCAGGTGCCACACCTCCCTGGATGATCCA 783  
Db 220 TCAGGATATACAGAGTTGGTAAATCACTCAGGTGCCACACCTCCCTGGATGATCCA 279  
Qy 784 AGATGAAGATACATTCGTGGGAACCAAGAATAGGACCATCTATGAAGATTTCTTAA 843  
Db 280 AGATGAAGATACATTCGTGGGAACCAAGAATAGGACCATCTATGAAGATTTCTTAA 339  
Qy 844 AGAAAGGAAACACAGAGTTGAAATACTCCCCCAGACCGAGTTGGGGCCAACTTTGA 903  
Db 340 AGAAAGGAAACACAGAGTTGAAATACTCCCCCAGACCGAGTTGGGGCCAACTTTGA 399  
Qy 904 TCACAGCTCCAGG 916  
Db 400 TCACAGCTCCAG 412  
RESULT 13  
LOCUS CQ727443 433 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 13377 from Patent WO02068579.  
ACCESSION CQ727443  
VERSION CQ727443.1 GI:42293443  
KEYWORDS Homo sapiens (human)  
SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 13377 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source Location/Qualifiers  
1..433  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 22.5%; Score 281.4; DB 6; Length 433;  
Best Local Similarity 98.7%; Pred. No. 2.2e-63;  
Matches 315; Conservative 0; Mismatches 1; Indels 3; Gaps 3;  
Qy 727 GGATATACCGAGCTTGTAAACATCCAC-TCAGGTGCCACACCTCCCTGGATGATCCAAG 785  
Db 117 GGATATACCGAGCTTGTAAACATCCAC-TCAGGTGCCACACCTCCCTGGATGATCCAAG 176  
Qy 786 ATGAAGAATACTTCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAG 845  
Db 177 ATGAAGAATACTTCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAG 236  
Qy 846 AAAAGGAAACAGAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATC 905  
Db 237 AAAAGGAAACAGAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATC 296  
Qy 906 ACAGCTCCAGGACGAGTGCAGGCTGCGCTCTTTTGGCCCGCTGGAATAATGGA 965  
Db 297 ACAGCTCCAGGACGAGTGCAGGCTGCGCTCTTTTGGCCCGCTGGAATAATGGA 355  
Qy 966 CGCCGCTGGAGTCCAGACATCAACTCCAAACTGGAAGCTCAGCAATGAAGAAGCAGTC 1025  
Db 356 CGCCGCTGGAGTCCAGACATCAATTCNAACTGGAAGCTCAGCAATGAAGAAGCAGTC 414  
Qy 1026 ACATACAGAAAAAGCTAA 1044  
Db 415 ACATACAGAAAAAGCTAA 433  
RESULT 14  
LOCUS CQ728296 263 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 14230 from Patent WO02068579.  
ACCESSION CQ728296  
VERSION CQ728296.1 GI:42296746  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 14230 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source Location/Qualifiers  
1..263  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 21.0%; Score 263; DB 6; Length 263;  
Best Local Similarity 100.0%; Pred. No. 1.6e-58;

Mon Jun 13 13:27:13 2005

Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 GTGAGGGGGCCGCAAGGCCATCCGCGCTCAGGTGGAGCGCTATGTCCCGAACAC 223  
Db 1 GTGGAGGGGGCCGCAAGGCCATCCGCGCTCAGGTGGAGCGCTATGTCCCGAACAC 60

QY 224 GAGCGATGCTGTGTGCTGTGCTGCGGCTGTGAGGTGCGGGAACACCTGAGCCATGGA 283  
Db 61 GAGCGATGCTGTGTGCTGTGCTGCGGCTGTGAGGTGCGGGAACACCTGAGCCATGGA 120

QY 284 AACCTGACGGTCTGTACGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAA 343  
Db 121 AACCTGACGGTCTGTACGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAA 180

QY 344 GCAACCAACAAATTCGTGGGAGAACAAAGCTGAGTCCAGATGAAAGAGAGTTTCTG 403  
Db 181 GCAACCAACAAATTCGTGGGAGAACAAAGCTGAGTCCAGATGAAAGAGAGTTTCTG 240

QY 404 GTCACTCCCAGGATTATGCGG 426  
Db 241 GTCACTCCCAGGATTATGCGG 263

RESULT 15  
AX906736 374 bp DNA linear PAT 18-DEC-2003  
LOCUS  
DEFINITION Sequence 22599 from Patent EP1033401.  
ACCESSION AX906736  
VERSION AX906736.1 GI:40061719  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Dumas Milne Edwards, J.B., Duclert, A. and Giordano, J.Y.  
TITLE Expressed sequence tags and encoded human proteins  
JOURNAL Patent: EP 1033401-A 22599 06-SEP-2000;  
Genet (FR)

FEATURES  
source 1..374  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 19.5%; Score 244.8; DB 6; Length 374;  
Best Local Similarity 98.5%; Pred. No. 1.1e-53;  
Matches 268; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 81 CCTTCTTCTGTGGTCGC-GGCACGTTTACAGCGCGCAAGCACCCAGCGGCGAGCTGAAGGAG 139  
Db 104 CCTTCTTCTGTGGTCGGCGGACGTTTACAGCGCGCAAGCA-CCAGCGGCGAGCTGAAGGAG 162

QY 140 GCTTTTGGAGAGGCTCTCTGCCCGAGTGGAGCGCGCCGCAAGGCCATCCCGCGCGCTCAG 199  
Db 163 GCTTTTGGAGAGGCTCTCTGCCCGAGTGGAGCGCGCCGCAAGGCCATCCCGCGCGCTCAG 222

QY 200 GTGAGCGCTATGTGCGCGAACACAGCGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGAG 259  
Db 223 GTGAGCGCTATGTGCGCGAACACAGCGGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGAG 282

QY 260 GTGCGGGAACACCTGAGCCATGGAACCTGACGGTGTGTACGGGGGGCTGTGAGGAT 319  
Db 283 GTGCGGGAACACCTGAGCCATGGAACCTGACGGTGTGTACGGGGGGCTGTGAGGAT 342

QY 320 CTGCGCGCGCGAGCGCAAGAACAAAGCAACCAA 351  
Db 343 CTGCGCGCGCGAGCGCAAGAACAAAGCAACCAA 374

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 425.87 Seconds  
(without alignments)  
17417.120 Million cell updates/sec

Title: US-09-155-676B-4  
Perfect score: 1253  
Sequence: 1 cattggagtcacgcggtggc.....ggggggccgtaccacctttt 1253

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870567 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	99.9	1253	2	AAV03325 Sequence
2	514.6	41.1	681	5	AA83767 DNA encod
3	479.6	38.3	2557	5	AA83767 DNA encod
4	451.8	36.1	3948	6	AA88468 Human leg
5	451.8	36.1	3948	10	AA88468 Human leg
6	431.8	36.1	3948	12	ADJ71904 Human lgs
7	400.8	32.0	2556	2	AAV03325 Human lgs
8	392	31.3	401	5	AA83767 Human sec
9	370.4	29.6	742	5	AA83766 DNA encod
10	300.4	24.0	2489	9	ADA45162 Human pol
11	297	23.7	421	2	AAV7867 EST clone
12	263.4	21.0	3100	5	AA83767 DNA encod
13	244.8	19.5	374	3	AA818524 Human sec
14	226	18.0	326	2	AA83767 DNA encod
15	222	17.7	584	10	ADB49603 Primary r
16	215.2	17.2	4017	5	AA83768 DNA encod
17	175.8	14.0	2803	4	AA81293 Human imm
18	175.8	14.0	3731	4	AA81295 Human imm
19	175.8	14.0	5269	4	AA81294 Human imm
20	171	13.6	695	4	AAK62424 Human imm

21	140.6	11.2	214	4	AAK61910 Human imm
22	132.8	10.6	1600	5	AA83768 DNA encod
23	130.8	10.4	467	5	AA83765 DNA encod
24	122.8	9.8	328	3	AA830122 Human sec
25	104.8	8.4	258	6	ABZ08765 Human leu
26	83.4	6.7	97	3	AA811553 Human sec
27	61.6	4.9	1906	2	AAV03323 5' end of
28	54.6	4.4	2000	8	ADA71938 Rice gene
29	53.2	4.2	2631	2	AAV03324 Clone 10
30	46.8	3.7	2126	2	AAV41999 Human GPR
31	46.8	3.7	2126	2	AAZ11174 Human GPR
32	46	3.7	10732	3	AA810594 Gene enco
33	45.8	3.7	864	11	ACN82266 Breast ca
34	45	3.6	647	10	ADD20719 Oreochrom
35	44.8	3.6	660	8	ABZ51586 Aspergill
36	44.8	3.6	801	8	ABZ51618 Aspergill
37	44.8	3.6	1236	2	AA806084 S. pneumo
38	44.8	3.6	1236	4	AAI70396 Streptoco
39	44.4	3.6	1236	10	ADC01891 S. pneumo
40	44.4	3.5	437	5	ABV31836 Human pro
41	43.8	3.5	1889	4	ABL57912 Rat VG41
42	43.8	3.5	1980	2	AA859454 H-Delta-1
43	43.2	3.4	468	5	ABV31846 Human pro
44	43.2	3.4	468	5	ABV40808 Human pro
45	43.2	3.4	478	5	ABV31844 Human pro

ALIGNMENTS

RESULT 1  
AAV03325  
ID AAV03325 standard; cDNA; 1253 BP.  
AC AAV03325;  
XX 15-APR-1998 (first entry)  
DE Sequence of clone 15, which encodes a TRAF2 binding protein.  
KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;  
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;  
KW intracellular signalling activity; acute hepatitis;  
KW autoimmune-induced cell death; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX CDS 2..1252  
XX FT /\*tag= a  
XX FT /note= "no ATG start or STOP codon given"  
XX  
XX WO9737016-A1.  
XX 09-OCT-1997.  
XX 01-APR-1997; 97WO-IL000117.  
XX 02-APR-1996; 96IL-00117800.  
XX 26-AUG-1996; 96IL-00119133.  
XX (YEDA ) YEDA RES & DEV CO LTD.  
XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;  
XX WPI; 1997-503101/46.  
XX P-PSDB; AAW42401.  
XX DNA encoding tumour necrosis factor receptor-associated factor binding  
XX molecule - used for modulation or mediation in cells of the activity of  
XX NF-KB..  
XX Claim 4; Fig 5; 127pp; English.

XX The present sequence is that of clone 15, which encodes a TRAF2 binding  
CC protein. Clone 15 is a partial clone, which lacks most of its 5' end of  
CC the coding DNA sequence. A cDNA library prepared from B-cells was  
CC screened for proteins that associate with TRAF2, and the present sequence  
CC isolated. The clone 15 protein is capable of binding to at least amino  
CC acids 222-501 of TRAF2. The TRAF-2 binding proteins can be used for  
CC modulation or mediation in cells of the activity of NF-kappaB or any  
CC other intracellular signalling activity modulated or mediated by TRAF2.  
CC TRAF-binding proteins are especially used for prevention or treatment of  
CC pathological conditions associated with NF-kB induction, e.g. acute  
CC hepatitis, autoimmune-induced cell death, e.g. death of the beta  
CC Langerhans cells or the pancreas that results in diabetes, the death of  
CC cells in graft rejection, the death of oligodendrocytes in the brain in  
CC multiple sclerosis, and AIDS-inhibited T cell suicide which causes  
CC proliferation of the AIDS virus and hence the AIDS disease. The proteins  
CC are also useful for screening of ligands capable of binding to a protein,  
CC which are useful for modulating cellular activity modulated/mediated by  
CC TRAF2  
XX

SQ Sequence 1253 BP; 341 A; 321 C; 334 G; 256 T; 0 U; 1 Other;

Query Match 99.9%; Score 1252; DB 2; Length 1253;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATTGGAGTCACGCGTGGCGGCTCTAGATAGTGGATCCCGGGCTGCANGGAATTC 60  
DB 1 CATTGGAGTCACGCGTGGCGGCTCTAGATAGTGGATCCCGGGCTGCANGGAATTC 60

QY 61 GATTCAGGCCAGCAAGGCCCTTCTCTGTGGTGGCGGCGCTTTACAGCCGCAAGCAC 120  
DB 61 GATTCAGGCCAGCAAGGCCCTTCTCTGTGGTGGCGGCGCTTTACAGCCGCAAGCAC 120

QY 121 CCAGCGGCAGCTCAAGGAGGCTTTTCAGAGGCTCTCTGCCCGGAGGCGCGCCGCAA 180  
DB 121 CCAGCGGCAGCTCAAGGAGGCTTTTCAGAGGCTCTCTGCCCGGAGGCGCGCCGCAA 180

QY 181 GGCCATCCGCGCGCTCAGTGGAGCGCTATGTGCGCCGCAACACAGCGATGCTGGTGG 240  
DB 181 GGCCATCCGCGCGCTCAGTGGAGCGCTATGTGCGCCGCAACACAGCGATGCTGGTGG 240

QY 241 CCTGTGCTGGCGCTGTGAGTGGCGGGAACACCTGAGCCATGGAAACCTGACGGTGTGTA 300  
DB 241 CCTGTGCTGGCGCTGTGAGTGGCGGGAACACCTGAGCCATGGAAACCTGACGGTGTGTA 300

QY 301 CGGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAAGCAACCAAAATTCG 360  
DB 301 CGGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCAAGAAAGCAACCAAAATTCG 360

QY 361 GTGGGGAACAAAGCTGAGTCCAGATGAAGAGAGTTTCTGTCACCTCCCGAGGATTA 420  
DB 361 GTGGGGAACAAAGCTGAGTCCAGATGAAGAGAGTTTCTGTCACCTCCCGAGGATTA 420

QY 421 TGCCCGGATTCAGAAATCCATGTGTAAGGTTTGGATTCTCTATCAAGAAAGGAGGATAA 480  
DB 421 TGCCCGGATTCAGAAATCCATGTGTAAGGTTTGGATTCTCTATCAAGAAAGGAGGATAA 480

QY 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGAGGTGGAGCAGCGCAGAGGTGGT 540  
DB 481 AGTGATCAAGGAGATGGCAGCTCAGATCCGAGGTGGAGCAGCGCAGAGGTGGT 540

QY 541 TCGGTCTGTCTAGAGCTCAGGCGATGCGCAGACCCAGAGAGGGCTCTTCAGCACCTAG 600  
DB 541 TCGGTCTGTCTAGAGCTCAGGCGATGCGCAGACCCAGAGAGGGCTCTTCAGCACCTAG 600

QY 601 AAGCTGGAAGGGATGAACAGCCAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGA 660  
DB 601 AAGCTGGAAGGGATGAACAGCCAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGA 660

QY 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCACTCTGACATTCATTGG 720  
DB 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCACTCTGACATTCATTGG 720

QY 721 CCATCAGGATATACAGGAGTTGGTAACATCCATCAGTGGCCACACCTCCCTCGGATGAT 780  
DB 721 CCATCAGGATATACAGGAGTTGGTAACATCCATCAGTGGCCACACCTCCCTCGGATGAT 780

QY 781 CCAAGATGAAGAATACATCTCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCT 840  
DB 781 CCAAGATGAAGAATACATCTCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCT 840

QY 841 TAAAGAAAAGGAAAAACAGAACTTGAAGAACTCCCGCCAGACCGAGTTGGGCGCAACTT 900  
DB 841 TAAAGAAAAGGAAAAACAGAACTTGAAGAACTCCCGCCAGACCGAGTTGGGCGCAACTT 900

QY 901 TGATCAGAGTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGTCTGGAATA 960  
DB 901 TGATCAGAGTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGTCTGGAATA 960

QY 961 ATGGACCCGCTGGCAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAAG 1020  
DB 961 ATGGACCCGCTGGCAGTCCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAAG 1020

QY 1021 CAGTCACATACAGAAAAGCTAATCATGCTCTACCAACTACCATGAGGCTAAAAGCC 1080  
DB 1021 CAGTCACATACAGAAAAGCTAATCATGCTCTCTCAACTACCATGAGGCTAAAAGCC 1080

QY 1081 AAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAATCTTTATCATTTGTTCTTAG 1140  
DB 1081 AAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAATCTTTATCATTTGTTCTTAG 1140

QY 1141 GAAACAGACATACATCTATTGATTAAGTTTATTTTCGGCCCTTCGTTGGCCT 1200  
DB 1141 GAAACAGACATACATCTATTGATTAAGTTTATTTTCGGCCCTTCGTTGGCCT 1200

QY 1201 CGAATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCGGTACCCACTTTT 1253  
DB 1201 CGAATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCGGTACCCACTTTT 1253

RESULT 2  
AAS83767  
ID AAS83767 standard; cDNA; 681 BP.  
XX  
AC AAS83767;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #19571.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSB-) HYSEQ INC.  
XX  
PI Dmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
XX  
PT P-PSDB; ABG19580.  
PT  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.

XX	PS	Claim 1; SEQ ID NO 19571; 103pp; English.
XX	XX	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal CC activity of (II) or to treat disease states involving (II). (II) is CC useful for generating antibodies against it, detecting or quantitating a CC polypeptide in tissue, as molecular weight markers and as a food CC supplement. (II) and its binding partners are useful in medical imaging CC involving aberrant protein expression or biological activity. The CC polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic CC coding sequences of the invention. Note: The sequence data for this CC patent did not appear in the printed specification, but was obtained in CC electronic format directly from WIPO at CC ftp.wipo.int/pub/published_pct_sequences
XX	SQ	Sequence 681 BP; 154 A; 171 C; 231 G; 125 T; 0 U; 0 Other;
		Query Match 41.1%; Score 514.6; DB 5; Length 681;
		Best Local Similarity 97.4%; Pred. No. 2.6e-140;
		Matches 568; Conservative 0; Mismatches 4; Indels 11; Gaps 4;
Qy	81	CTTCTTTCTGTGTCGC-GGCACGCTTTACAGCCGCAAGCACCCAGCGAGCTGAAGGAG 139
Db	99	CTTCTTTCTGTGTCGCGGCACGCTTTACAGCCGCAAGCA-CCAGCGCGAGCTGAAGGAG 157
Qy	140	GCTTTTGAGAGGCTCTGCCCC-----AGTGTGAGGCGCGCCCGCAAGCCATCCGCG 191
Db	158	GCTTTTGAGAGGCTCTGCCCCAGGTGCGGAGGTGGAGGCGCGCCCGCAAGCCATCCGCG 217
Qy	192	CCGCTCAGGTGGAGCGCTATGTGCCGAACACGAGCGATGCTGCTGTGCTGTGCTGCG 251
Db	218	CCGCTCAGGTGGAGCGCTATGTGCCGAACACGAGCGATGCTGCTGTGCTGTGCTGCG 277
Qy	252	GCTGTGAGGTGCGGGAACACCTGAGCCATGAAACCTGACGGTCTCTAGCGGGGGCTGC 311
Db	278	GCTGTGAGGTGCGGGAACACCTGAGCCATGAAACCTGACGGTCTGTACGGGGGCTGC 337
Qy	312	TGGAGCATCTGGCCAGCCCGAGAGCACAAAGAACCAACAATTCGTGTGGGAGAACA 371
Db	338	TGGAGCATCTGGCCAGCCCGAGAGCACAAAGAACCAACAATTCGTGTGGGAGAACA 397
Qy	372	AAGCTGAGGTCCAGATGAAAGAGAGTTCTGTGTCACCTCCCGAGGATTATGCGGATTCA 431
Db	398	AAGCTGAGGTCCAGATGAAAGAGAGTTCTGTGTCACCTCCCGAGGATTATGCGGATTCA 457
Qy	432	AGAAATCCATGGTGAAGGTTTGGATTTCCTATGAAGAAAAGGAGGATAAAGTGCATCAGG 491
Db	458	AGAAATCCATGGTGAAGGTTTGGATTTCCTATGAAGAAAAGGAGGATAAAGTGCATCAGG 517
Qy	492	AGATGGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACA-GGAGGTGGTTGGTCTGTCTC 550
Db	518	AGATGGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAAGGAGGTGGTTGGTCTGTCTC 577
Qy	551	TTAGAGCTCAGGAGTGCAGAGCCAGAGAGGGCTTTTCAGCACTTAGAGCTGAGAA 610
Db	578	TTAGAGCTCAGGAGTGCAGAGCCAGAGAGGGCTTTTCAGCACTTAGAGCTGAGAA 637
Qy	611	GGGATGAACGCCAAGTAGCTTTCAGCTTACAGCAGCCCTCAA 653
Db	638	GGGATGAACGCCAAGTAGCTTTCAGCTTACAGCAGCCCTTAA 680
		RESULT 3

AAS87242	AA887242 standard; cDNA; 2557 BP.	
XX	AC	AAS87242;
XX	DT	13-FEB-2002 (first entry)
XX	DE	DNA encoding novel human diagnostic protein #23046.
XX	KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	OS	Homo sapiens.
XX	PN	WO200175067-A2.
XX	PD	11-OCT-2001.
XX	PF	30-MAR-2001; 2001WO-US008631.
XX	PR	31-MAR-2000; 2000US-00540217.
XX	PR	23-AUG-2000; 2000US-00649167.
XX	PA	(HYSE-) HYSEQ INC.
XX	PI	Drmanac RT, Liu C, Tang YT;
XX	DR	WPI; 2001-639362/73.
XX	DR	P-PSDB; ABG23055.
XX	PT	New isolated polynucleotide and encoded polypeptides, useful in
XX	PT	diagnostics, forensics, gene mapping, identification of mutations
XX	PT	responsible for genetic disorders or other traits and to assess
XX	PT	biodiversity.
XX	PS	Claim 1; SEQ ID NO 23046; 103pp; English.
XX	CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX	CC	sequences. (I) is useful as hybridisation probes, polymerase chain
XX	CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX	CC	and in recombinant production of (II). The polynucleotides are also used
XX	CC	in diagnostics as expressed sequence tags for identifying expressed
XX	CC	genes. (I) is useful in gene therapy techniques to restore normal
XX	CC	activity of (II) or to treat disease states involving (II). (II) is
XX	CC	useful for generating antibodies against it, detecting or quantitating a
XX	CC	polypeptide in tissue, as molecular weight markers and as a food
XX	CC	supplement. (II) and its binding partners are useful in medical imaging
XX	CC	of sites expressing (II). (I) and (II) are useful for treating disorders
XX	CC	involving aberrant protein expression or biological activity. The
XX	CC	polypeptide and polynucleotide sequences have applications in
XX	CC	diagnostics, forensics, gene mapping, identification of mutations
XX	CC	responsible for genetic disorders or other traits to assess biodiversity
XX	CC	and to produce other types of data and products dependent on DNA and
XX	CC	amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic
XX	CC	coding sequences of the invention. Note: The sequence data for this
XX	CC	patent did not appear in the printed specification, but was obtained in
XX	CC	electronic format directly from WIPO at
XX	CC	ftp.wipo.int/pub/published_pct_sequences
XX	SQ	Sequence 2557 BP; 635 A; 624 C; 691 G; 606 T; 0 U; 1 Other;
		Query Match 38.3%; Score 479.6; DB 5; Length 2557;
		Best Local Similarity 84.9%; Pred. No. 1.1e-129;
		Matches 702; Conservative 0; Mismatches 55; Indels 70; Gaps 12;
Qy	426	GATTCAAGAAATCCATGGTGAAGGTTTGGATTCCCTATGAAGAAAAGGAGGATAAAGTGA 485
Db	36	GATTCAAGAAATCCATGGTGAAGGTTTGGATTCCCTATGAAGAAAAGGAGGATAAAGTGA 95
Qy	486	TCAGGAGTGGCAGCTCAGATCCGTGAGGTGGAGCAGCCGACAGAGGTGCTCGGT 545
Db	96	TCAGGAGTGGCAGCTCAGATCCGTGAGGTGGAGCAGCCGACAGAGGTGCTCGGT 155

QY	546	CTGTCTTAGAG-----CCTCAGGCACTGCCAGACCCA	577
Db	156	CTGTCTTAGAGGTTGGTTCCCTCGGAGGATCCAGACCACTCAGGCACTGCCAGACCCA	215
QY	578	GAAGAGGGCTCTTCAGCACTTA-GAAGCTGGAAAGGATGAACAGCCAAAGTAGCTT-CCA	635
Db	216	GAAGAGGGCTCTTCAGCACTTAGGAAGCTGGAAAGGATGAACAGCCAAAGTAGCTTCCA	275
QY	636	GCTTACAGCAGCCCTCAAAATTTGGACCTGCCACACAGCTCCAGAGCTTGACTGGATGGAGA	695
Db	276	GCTTACAGCAGCCCTCAAAATTTGGACCTGCCACACAGCTCCAGAGCTTGACTGGATGGAGA	335
QY	696	CAGGACCATCTCTGACATTCATTGGCCATCAGATATA-CAAGGAGTTGGTAAACATCCAC	754
Db	336	CAGGACCATCTCTGACATTCATTGGCCATCAGATATATCCAGGAGTTGGTAAACATCCAC	395
QY	755	TCAGGTGCCACACTCCCTCGGATGATCCAGATGAAGAATACAT----TGCTGGGAACA	810
Db	396	TCAGGTGCCACACTCCCTCGGATGATCCAGATGAAGAATACATTTGGCTGGGGGACCCCA	455
QY	811	AGAAATAGGACC---ATCCTATGAAGATTTCTTAAAGAAA-----	848
Db	456	AGAAATNGGGCCCCATCCCTATGAAGATTTCTTAAAGAAATTTCCCTATCATGTATAC	515
QY	849	--AGGAAAAACAGAGTTGAAAAAACTCCCCC-----AGACCGAGTTGGGG--CCAAC	900
Db	516	AGAGGAAAAACAGAGTTGAAAAAACTCCCCC-----AGACCGAGTTGGGG--CCAAC	575
QY	901	TGATCACAGCTCC--AGGACCAAGTGAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAA	958
Db	576	GGATCACAGCTCCAGGACCCAGTGCGAGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAA	635
QY	959	TAATGACCGCGCTGCGAGTCCAGACATCAATCCAAACTGAAGTGCAGCAATGAAGA	1018
Db	636	AATCGGACCGCGTGCGAGTCCAGACATCAATCCAAACTGAAGTGCAGCAATGAAGA	695
QY	1019	AGCAGTCAATACAGAAAAAGCTAATCATGTCTCTACCAACTACCATGAGGCTAAAAG	1078
Db	696	AGCAGTCAATACAGAAAAAGCTAATCATGTCTCTACCAACTACCATGAGGCTAAAAG	755
QY	1079	CCAAAGTCAACCAACCCCTATATATACCTTCCACCCAAATTTCTTATCATGTCTTTCTT	1138
Db	756	C--AAAGTCAACAAACCCCTATATATACCTTCCA-CCAAATTTCTTATCATGTCTTTCTT	812
QY	1139	AGGAAACAGACATACATCATTTGATTTAATAAAGTTTATTTT 1185	
Db	813	AGGAAACAGACATACATCATTTGATTTAATAAAGTTTAAATTT 859	
RESULT 4			
ID	AAF88468	standard; cDNA; 3948 BP.	
AC	AAF88468;		
XX			
DT	18-NOV-2002	(first entry)	
XX			
DE	Human legless homologue hlgs-1 cDNA.		
XX			
KW	Legless; human; lgs; Wnt/Wingless signaling pathway; Wnt; Wg; gene;		
KW	tissue proliferation; tumour; cytostatic; cellular disorder; colon;		
KW	blood disorder; cancer; breast; head and neck cancer; brain; thyroid;		
KW	medulloblastoma; skin cancer; tissue regeneration; tissue repair; ss.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	1..3948		
FT	CD5		
FT	/*tag= a		
FT	/product= "hlgs-1"		
FT	misc_feature		
FT	601..3948		
FT	/*tag= b		
FT	/note= "region encoding protein represented in protein		



[illegible]





XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 742 BP; 248 A; 180 C; 145 G; 169 T; 0 U; 0 Other;  
Query Match 29.6%; Score 370.4; DB 5; Length 742;  
Best Local Similarity 97.4%; Pred. No. 7e-98;  
Matches 419; Conservative 0; Mismatches 6; Indels 5; Gaps 4;  
QY 757 AGTGCCACACCTCCCTGGATGATCCAGATGAAGTAATACATTGCTGGGAACCAAGAAAT 816  
Db 306 AGGTGCCACACCTCCCTGGATGATCCAGATGAAGTAATACATTGCTGGGAACCAAGAAAT 365  
QY 817 AGGACCATCTATGAGATTTCTTAAAGAAAAGGAAAACAGAGTTGAAAAAACTCCC 876  
Db 366 AGGACCATCTATGAGATTTCTTAAAGAAAAGGAAAACAGAGTTGAAAAAACTCCC 425  
QY 877 CCAGACCGAGTTGGGGCCAACTTTTGATCACAGCTCCAGGACGAGTGGCGTGGCTGCC 936  
Db 426 CCAGACCGAGTTGGGGCCAACTTTTGATCACAGCTCCAGGACGAGTGGCGTGGCTGCC 485  
QY 937 CTCCTTTTGGCGCGTCTGGATTAATGACGCGCTGGCGAGTCCAGACATCAACTCCAAA 996  
Db 486 CTCCTTTT-GGCCGCGTCTGGATTAATGACGCGCTGGCGAGTCCAGACATCAA-TTCAA 543  
QY 997 ACTGAAGCTGCAGCAATGAAGAGCAGTCCACATCAGAAAAGCTAATCATGCTCTTA 1056  
Db 544 ACTGAAGCTGCAGCAATGAAGAGCAGTCCACATCAGAAAAGCTAATCATGCTCTTA 603  
QY 1057 CCAACTACCATGAGGCTAAAGCCAAAGTCAACCAAAACCCCTATTATATACCTTCCACCCAA 1116  
Db 604 CCAACTACCATGAGGCTAAAGC--AAAGTCAACAAACCCCTATTATACCTTCCA-CCAA 660  
QY 1117 ATTCCTTATCATGCTCTTCTTAGAAGAACAGACATCACTCATTTGATTAATAAGT 1176  
Db 661 ATTCCTTATCATGCTCTTCTTAGAAGAACAGACATCACTCATTTGATTAATAAGT 720  
QY 1177 TTTATTTTTC 1186  
Db 721 TTTATTTTTC 730  
RESULT 10  
ID ADA45162  
XX ADA45162 standard; cDNA; 2489 BP.  
AC ADA45162;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human polynucleotide #72.  
XX  
KW Human; gene; ss; genetic disorder; genetic fingerprinting;

KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;  
KW insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;  
KW periodontal disease; bone fracture; cartilage damage;  
KW central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; cancer; nutrition; carbon source; nitrogen source;  
KW carbohydrate source.  
XX  
OS Homo sapiens.  
XX  
PN US2003044935-A1.  
XX  
XX 06-MAR-2003.  
PD  
XX 21-DEC-2000; 2000US-00746783.  
PF  
XX 11-JUN-1997; 97US-0086236P.  
PR 12-JUN-1997; 97US-0086234P.  
PR 08-JUL-1997; 97US-0092115P.  
PR 08-SEP-1997; 97US-0093045P.  
PR 27-OCT-1997; 97US-0090100P.  
PR 02-OCT-1997; 97US-00958304.  
PR 07-NOV-1997; 97US-0090111P.  
PR 05-JUN-1998; 98US-00092722.  
PR 11-JUN-1998; 98US-00096287.  
PR 17-JUN-1998; 98US-00098588.  
PR 04-AUG-1998; 98US-00130189.  
PR 08-SEP-1998; 98US-00149633.  
PR 01-OCT-1998; 98US-00165960.  
PR 04-NOV-1998; 98US-00185936.  
XX  
PA (JACO/) JACOBS K.  
PA (MCCO/) MCCOY J M.  
PA (LVALL/) LA VALLIE E R.  
PA (COLL/) COLLINS-RACIE L A.  
PA (EVAN/) EVANS C.  
PA (MERB/) MERBERG D.  
PA (TREA/) TREACY M.  
PA (SPAU/) SPAULDING V.  
XX  
XX Jacobs K, McCoy JM, La Vallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Spaulding V;  
PI  
XX WPI; 2003-521754/49.  
XX P-PSDB; ADA45163.  
XX  
XX New polypeptides and polynucleotides having biological activities, useful  
as nutritional sources or supplements, or for treating e.g. autoimmune  
diseases, cancers, bone fractures or damages, or central nervous system  
disorders.  
XX  
XX Disclosure; Page 237-238; 288pp; English.

The invention relates to human polynucleotides and the polypeptides they  
encode. The polynucleotides can be used to express recombinant proteins  
for analysis, characterisation or therapeutic use, as markers for tissues  
in which the corresponding protein is expressed, as molecular weight  
markers on Southern gels, as chromosome markers or tags to identify  
chromosomes or to map related gene positions, to compare with endogenous  
DNA sequences in patients to identify potential genetic disorders, as  
probes to hybridise and discover novel related DNA sequences, as a source  
of information to derive PCR primers for genetic fingerprinting, to raise  
anti-protein antibodies and in gene therapy. The proteins can be used to  
raise antibodies or to elicit another immune response, as reagents in  
assays designed to quantitatively determine levels of the protein in  
biological fluids, as markers for tissues in which the corresponding  
protein is preferentially expressed and to treat autoimmune disorders  
(e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent  
diabetes mellitus or graft-versus-host disease), anaemias, periodontal  
diseases, bone fractures, cartilage damage, central nervous system  
disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.  
The proteins and polynucleotides are also useful as nutritional sources  
or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This  
sequence represents a human polynucleotide of the invention.

```
XX SQ Sequence 2489 BP; 699 A; 556 C; 506 G; 723 T; 0 U; 5 Other;
Query Match 24.0%; Score 300.4; DB 9; Length 2489;
Best Local Similarity 91.5%; Pred. No. 5.4e-77;
Matches 432; Conservative 0; Mismatches 21; Indels 19; Gaps 10;

Qy 715 CATTGGCCATCAGATATACAGGAGTTGGTAAATCCATCCAGTGCACACCTCCCTG 774
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1969 CATTTCACCTACAGGATATAGGAGTTGGTAAATCCATCCAGTGCACACCTCCCTG 2028
Qy 775 GATGATCAAGATGAAGATACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGA 834
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2029 GATGATCAAGATGAAGA--ACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGA 2086
Qy 835 ATTTCTTAAAGAAAGAAACAGAGTTGAAAGAACTCCCCCGACAGCTGGGGC 894
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2087 ATTTCTTAAAGAAAGG--AAACAGAGTTGAAAGAACTCCCCCGACAGCTGGGGC 2144
Qy 895 CAACCTTGATCAGAGTCCAGGACAGTGCAGGCTGGCTGCTCTTTTGGGCGCGTCT 954
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2145 CAACCTTGATCAGAGTCC--GACCAGTGCAGGCTGGCTGCTCTTTT-GGCGGTGCT 2201
Qy 955 GGAATAATGGACGCGCTGGCAGTCCAGACATCACTCCAAACTGAAGCTGCAGCAATG 1014
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2202 GGAATAATGGACGCGCTGG--GTCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATG 2258
Qy 1015 AAGAAGCAGTCAATACAGAAAGAGTAAATCATGCTCTCTPACCAACTACCATGAGGCTA 1074
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2259 AAGAAGCAGTCAATACAG--AAAAGCTAATCATGCTCTCTACCACTACCATGAGGCTA 2316
Qy 1075 AAAGCCAAAGTCAACCAACCCCTATTATATACCTTCCACCCCAAAATCTTTATCATGCTT 1134
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2317 AAAG-CAAAGTCAACCAACCCCTTAT---ACCTTCCACCAAAATCTTTATCATGCTT 2371
Qy 1135 TCTTAGGAAACAGACATCACTCATTTGATTTTAAATAAGTTTTATTTTC 1186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2372 TCTTAGGAAACAGACATCACTCATTTGATTTTAAATAAGTTTTATTTTC 2421

RESULT 11
AAV86432
ID AAV86432 standard; cDNA; 421 BP.
AC AAV86432;
AT 27-APR-1999 (first entry)
DE EST clone AS20.
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemoraxis; chemokines; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
OS Homo sapiens.
XX
XX WO9845435-A2.
XX
XX 15-OCT-1998.
XX
XX 10-APR-1998; 98WO-US006954.
XX
XX 10-APR-1997; 97US-00835913.
XX
XX (GEMV ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
XX Spaulding V, Agostino MJ;
XX WPI; 1999-070076/06.
XX
XX New polynucleotides encoding human secreted proteins - derived from e.g.
```

```
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT pituitary, retina and colon cDNA libraries.
XX Claim 1; Page 232; 633pp; English.
XX
XX This sequence represents an expressed sequence tag (EST), and is a
XX polynucleotide of the invention. The polynucleotides of the invention are
XX all secreted EST sequences isolated from a variety of human tissue
XX sources. The EST sequences and proteins encoded by them are predicted to
XX have useful biological activities which would make them suitable for
XX treating, preventing or ameliorating medical conditions in humans and
XX animals, although no supporting data is given. Suggested activities
XX include nutritional activity, immune stimulating or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX activity. The EST sequences are also stated to be useful for gene therapy
XX
XX Sequence 421 BP; 128 A; 110 C; 88 G; 95 T; 0 U; 0 Other;
Query Match 23.7%; Score 297; DB 2; Length 421;
Best Local Similarity 96.8%; Pred. No. 2e-76;
Matches 303; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 604 CTGGAAGGATGAACAGCCAAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTTGGACCT 663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 CTGCTCATGCTTTTTCAGCCAAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTTGGACCT 159
Qy 664 GCCACCAAGTCCAGAGCTTCACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCA 723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
160 GCCACCAAGTCCAGAGCTTCACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCA 219
Qy 724 TCAGGATATACAGGAGTGGTAAACATCCACTCAGGTGCCACACCTCCTCGATGATCCA 783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
220 TCAGGATATACAGGAGTGGTAAACATCCACTCAGGTGCCACACCTCCTCGATGATCCA 279
Qy 784 AGATGAAGAATACATTCTCGGGAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAA 843
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
280 AGATGAAGAATACATTCTCGGGAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAA 339
Qy 844 AGAAGGAAAGAAACAGAGTTGAAAGAACTCCCGCCAGACCGAGTTGGGGCCACTTTGA 903
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 AGAAGGAAAGAAACAGAGTTGAAAGAACTCCCGCCAGACCGAGTTGGGGCCACTTTGA 399
Qy 904 TCACAGCTCCAGG 916
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 TCACAGCTCCAG 412

RESULT 12
AAS77867
ID AAS77867 standard; cDNA; 3100 BP.
AC AAS77867;
AT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #13671.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
```

```
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX PR WPI; 2001-639362/73.
XX DR P-PSDB; ABG13680.
XX XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID NO 13671; 103pp; English.
XX XX
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX CC coding sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 3100 BP; 759 A; 849 C; 927 G; 564 T; 0 U; 1 Other;
XX XX
Query Match 21.0%; Score 263.4; DB 5; Length 3100;
Best Local Similarity 99.6%; Pred. No. 4.7e-66;
Matches 264; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 162 AGGTGAGGCGCGCCGCAAGGCCATCCGCGCGCTCAGGTGAGGCGCTATGTGCCCGAAC 221
Db ATGTGAGGCGCGCGCCGCAAGGCCATCCGCGCGCTCAGGTGAGGCGCTATGTGCCCGAAC 448
QY 222 ACGAGCGATGCTGTGGTGCCTGTGCTGCGGCTGTGAGTGCGGGAAACCTGAGCCATG 281
Db ACGAGCGATGCTGTGGTGCCTGTGCTGCGGCTGTGAGTGCGGGAAACCTGAGCCATG 508
QY 282 GAAACCTGACGCTGTGTACGGGGGGCTGTGAGCAGCTGTGAGCAGCCAGGACACAGA 341
Db GAACCTGACGCTGTGTACGGGGGGCTGTGAGCAGCTGTGAGCAGCCAGGACACAGA 568
QY 342 AAGCAACCAACAAATTTCTGGTGGGAGAACAAAGCTCAGGTCCAGATGAAAGAGAGTTTC 401
Db AAGCAACCAACAAATTTCTGGTGGGAGAACAAAGCTCAGGTCCAGATGAAAGAGAGTTTC 628
QY 402 TGGTCACCTCCCGAGGATATGCGCG 426
Db TGGTCACCTCCCGAGGATATGCGCG 653
RESULT 13
AAC18524
ID AAC18524 standard; cDNA; 374 BP.
XX AAC18524;
XX AC
XX DT 06-OCT-2000 (first entry)
XX DE Human secreted protein 5' EST, SEQ ID NO: 22599.
```

```
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS Homo sapiens.
XX PN EP1033401-A2.
XX PD 06-SEP-2000.
XX PF 21-FEB-2000; 2000EP-00200610.
XX PR 26-FEB-1999; 99US-0122487P.
XX XX (GEST ) GENSET.
XX PA Dumas Milne Edwards J, Duclert A, Giordano J;
XX PI WPI; 2000-500381/45.
XX DR
XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX PS Claim 1; SEQ ID NO 22599; 71pp + Sequence Listing; English.
XX XX
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
XX CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
XX CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
XX CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors
XX SQ Sequence 374 BP; 66 A; 109 C; 139 G; 59 T; 0 U; 1 Other;
XX XX
Query Match 19.5%; Score 244.8; DB 3; Length 374;
Best Local Similarity 98.5%; Pred. No. 4.1e-61;
Matches 268; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 81 CCTTCTTCTGTGTGTCGC-GGCACGTTTTACAGCCGCAAGCACCCAGCGGACGCTGAAGGAG 139
Db 104 CCTTCTTCTGTGTGTCGGGGCACGTTTACAGCCGCAAGCA-CCAGCGGACGCTGAAGGAG 162
QY 140 GCTTTTGAGAGGCTCTGCCCCAGGTGAGGCGCGCCGCAAGGCCATCCGCGCGCTCAG 199
Db 163 GCTTTTGAGAGGCTCTGCCCCAGGTGAGGCGCGCCGCAAGGCCATCCGCGCGCTCAG 222
QY 200 GTGAGAGCGCTATGTGCCGAACACAGCGATGCTGTGGTCCCTGTGCTGGGCTGTGAG 259
Db 223 GTGAGAGCGCTATGTGCCGAACACAGCGATGCTGTGGTCCCTGTGCTGGGCTGTGAG 282
QY 260 GTGCGGGAACACCTTGAGCCCATGGAACCTGACGGTGTCTGTACGGGGGGCTGCTGGAGCAT 319
Db 283 GTGCGGGAACACCTTGAGCCATGGAACCTGACGGTGTCTGTACGGGGGGCTGCTGGAGCAT 342
QY 320 CTGGCCAGCCGAGAGCAAGCAAGCAACCA 351
Db 343 CTGGCCAGCCGAGAGCAAGCAAGCAACCA 374
RESULT 14
AAT24301
ID AAT24301 standard; cDNA to mRNA; 326 BP.
XX AAT24301;
XX AC
```



Query Match		17.7%;	Score 222;	DB 10;	Length 584;
Best Local Similarity		76.0%;	Pred. No. 2.7e-54;		
Matches 327;		Conservative 0;	Mismatches 95;	Indels 8;	Gaps 4;
QY	757	AGGTGCCACACCTCCCTGGATGATCCAAGATGAAGAATACATTTGCTGGGAACCAAGAAAT	816		
Db	423	AGGTGCCACACCTCCATGGATGATCCAGAGGAGGAGCACAGCTCTGGAAGCTCCCAAT	364		
QY	817	AGGACCATCCTATGAAGAATTTCTTAAAGAAAAGGAAAAACAGAAAGTTGAAAAAACTCCC	876		
Db	363	AGGACCTCCTATGAAGAATTTCTTAAAGAAAAGGAAAAACGAAAACTGAAGAACTCCC	304		
QY	877	CCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAGTCAGGCTGGCTGCC	936		
Db	303	TCCAGATAGAGTTGGGGCCAACTTTGATCACAGTTCCAACACCAAGTCAGGCTGGCTGCC	244		
QY	937	CTCTTTTGGGCGCGTCTGGAATATGGACGCGCTGGCAGTCCAGACATCAACTCCAAA	996		
Db	243	CTCTTTT-GGCAGAGTCTGGAATATGGACGCGCTGGCAGTCCAGGCATCAA-TTCAAA	186		
QY	997	ACTGAAGCTGCAGCAATGAAGAAGCAGTCACATACAGAAAAAGCTAATCATGCTCTCTA	1056		
Db	185	ACTGAAGCTGCAGCAAGAGCAGGCCACATATAAGGGAAAAAGCTGAAGTTCCCAATC	126		
QY	1057	CCAACTACCATGAGGCTAAAGCCAAAGTCAACCAACCCCTATTATACCTTCCACCCAA	1116		
Db	125	TGTTCCGATCTCATGGGTAA-CAACCCATTCAGCCAGCTATTATTCCTAAGACCAA	67		
QY	1117	ATTCTTTATCATTTGCTCTTTCTTAGGAAACAGACATACATTCATTGATTTTAATAAAGT	1176		
Db	66	-----TCCTCTCCAGCTGCCTGAGGAAACAGATGCACCCATTCATTGATTCAATAAAGT	12		
QY	1177	TTTATTTTTC	1186		
Db	11	TTTATTTTCC	2		

Search completed: June 10, 2005, 06:10:33  
Job time : 431.87 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 2640.16 Seconds  
(without alignments)  
18065.048 Million cell updates/sec

Title: US-09-155-676B-4  
Perfect score: 1253  
Sequence: 1 cattggagtcacgcggtggc.....ggggggccgtaccacactttt 1253

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	755.6	60.3	926	2	BE796421 601592158
2	713.2	56.9	884	4	BI859740 603387142
3	694	55.4	1065	4	BQ403663 602419385
C 4	616.8	49.2	714	5	BQ008160 UI-H-E11-
5	587	46.8	816	4	BI769735 603055036
C 6	578.4	46.2	652	5	BQ007036 UI-H-E11-
C 7	576.4	46.0	758	5	BQ007036 UI-H-E11-
C 8	572	45.7	760	4	BM682005 UI-E-E01-
9	565	45.1	878	4	BI414452 602986067
10	563.8	45.0	652	7	CH291332 170006000
C 11	558.4	44.6	743	5	BQ608980 UI-CF-FN0
C 12	558.4	44.6	747	5	BQ608985 UI-CF-FN0
C 13	548.6	43.8	656	4	BM684305 UI-E-EJ1-
14	541.4	43.2	596	5	BM929870 UI-E-EJ1-
C 15	531.8	42.4	668	4	BI046266 MR3-FN020
C 16	530.8	42.4	870	3	AK052757 Mus muscu
C 17	519.4	41.5	599	6	CB321979 UI-CF-FN0
C 18	509	40.6	629	6	CA438500 UI-H-DT1-
C 19	507.4	40.5	706	5	BQ006283 UI-H-E11-
C 20	495.4	39.5	634	6	CA748930 UI-H-FE1-
C 21	486.4	38.8	680	5	BU733365 UI-E-CQ1-
C 22	486.4	38.8	759	5	BU733404 UI-E-CQ1-
C 23	486.2	38.7	790	6	CB989556 AGENCOURT
C 24	484.4	38.7	635	6	CA424880 UI-H-FE1-

C 25	456.8	36.5	520	2	BF446390
26	456	36.4	635	9	AY402903
27	448.4	35.8	548	2	BE669606
28	445.4	35.5	468	2	BE504243
29	442.4	35.3	635	9	AY402904
30	442.2	35.3	666	2	BE503633
31	437.4	34.9	704	7	CR454777
32	427	34.1	493	5	BE281470
33	426.2	34.0	571	2	AW957079
C 34	424.2	33.9	501	5	BM976229
C 35	420.6	33.6	534	2	AW243066
36	413.6	33.0	976	2	BF531630
37	403	32.2	735	7	CN792842
38	402.4	32.1	538	2	AW014772
39	400.8	32.0	502	2	BE222549
40	396	31.6	485	1	AV666223
41	394.2	31.5	748	4	BG701634
C 42	390.4	31.2	452	1	AI697110
C 43	382.8	30.6	444	4	BG231618
44	382.4	30.5	765	7	CO798321
45	380.2	30.3	497	2	BF077048

ALIGNMENTS

RESULT 1  
BE796421  
LOCUS BE796421 926 bp mRNA linear EST 20-SEP-2000  
DEFINITION 601592158F1 NIH\_MGC\_7 Homo sapiens CDNA clone IMAGE:3945950 5', mRNA sequence.  
ACCESSION BE796421  
VERSION BE796421.1 GI:10217619  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 926)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-i@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: L1CM805 row: d column: 15  
High quality sequence stop: 784.  
FEATURES  
source  
Location/Qualifiers  
1. 926  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3945950"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/vector="pUC19"  
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 60.3%; Score 755.6; DB 2; Length 926;

Best Local Similarity 96.9%; Pred. No. 1e-195;		884 bp mRNA linear EST 10-OCT-2001	
Matches 814; Conservative 0; Mismatches 19; Indels 7; Gaps 4;		603387142P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:539596 5',	
		mRNA sequence.	
QY	81 CCTTCTTCTGTGTCGC-GGCACGTTTACAGCGCAACCCAGCGCAGCTGAAGAG	139	
Db	16 CTTTCTTCTGTGTCGCGGCACGTTTACAGCGCAAGCA-CCAGCGCAGCTGAAGAG	74	
QY	140 GTTTGTAGAGGCTCTGTCGCCAGGTGAGGGGGCCGCAAGGCCATCCGGCGGCTCAG	199	
Db	75 GCTTTGGAGAGGCTCTGTCGCCAGGTGAGGGGGCCGCAAGGCCATCCGGCGGCTCAG	134	
QY	200 GTGAGCGCTATGTCGCCGAACACAGAGGATGCTCTGCTGTGCTGTGCTGCGGCTGTGAG	259	
Db	135 GTGAGCGCTATGTCGCCGAACACAGAGGATGCTCTGCTGTGCTGTGCTGCGGCTGTGAG	194	
QY	260 GTGCGGGAACCTGAGCCATGGAACCTGACGCTGTGTCGCGGGGCTGCTGAGGAT	319	
Db	195 GTGCGGGAACCTGAGCCATGGAACCTGACGCTGTGTCGCGGGGCTGCTGAGGAT	254	
QY	320 CTGGCCAGCCAGAGCAAGAAGCAACCAAAATTTCTGTGGGAGCAAAAGCTGAG	379	
Db	255 CTGGCCAGCCAGAGCAAGAAGCAACCAAAATTTCTGTGGGAGCAAAAGCTGAG	314	
QY	380 GTCCAGATGAAGAGAGTTTCTGTGTCACCTCCCGAGGATTAATGCGGATTTCAAGAAATCC	439	
Db	315 GTCCAGATGAAGAGAGTTTCTGTGTCACCTCCCGAGGATTAATGCGGATTTCAAGAAATCC	374	
QY	440 ATGGTGAAGGTTTGGATTCTTATGAAGAAAGAGGATTAATGATCAAGAGATGGA	499	
Db	375 ATGGTGAAGGTTTGGATTCTTATGAAGAAAGAGGATTAATGATCAAGAGATGGA	434	
QY	500 GCTCAGATCGTAGGTGAGCAGAGCCGACAGGAGTGTGTCGCTCTTTAGAGCT	559	
Db	435 GCTCAGATCGTAGGTGAGCAGAGCCGACAGGAGTGTGTCGCTCTTTAGAGCT	494	
QY	560 CAGGAGTGCAGACCCAGAGAGGCTCTTTAGCACCCTTAGAAGCTGGAAGGATGAAC	619	
Db	495 CAGGAGTGCAGACCCAGAGAGGCTCTTTAGCACCCTTAGAAGCTGGAAGGATGAAC	554	
QY	620 AGCCAGTAGCTTCCAGCTTACAGAGCCCTCAATTTGGACCTGCCACAGCTCCAGAG	679	
Db	555 AGCCAGTAGCTTCCAGCTTACAGAGCCCTCAATTTGGACCTGCCACAGCTCCAGAG	614	
QY	680 CTTGACTGGATGAGCAGAGCACCCTCTGACATTCATTGGCCATCAGGATATACAGGA	739	
Db	615 CTTGACTGGATGAGCAGAGCACCCTCTGACATTCATTGGCCATCAGGATATACAGGA	674	
QY	740 GTTGGTAACATCCACTCAGGTGCCACACCTTCTGATGATCCAAAGATGAAGATACATT	799	
Db	675 GTTGGTAACATCCACTCAGGTGCCACACCTTCTGATGATCCAAAGATGAAGATACATT	734	
QY	800 GCTGGGAACCAAGAAATAGACCATCTATGAAGATTTCTTAAGAAAGGAAACAG	859	
Db	735 GCTGGGAACCAAGAAATAGACCATCTATGAAGATTTCTTAAGAAAGGAAACAG	794	
QY	860 AAGTTGAAAACCTCCCGCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACC	919	
Db	795 AGGTGAAGAAA---TCCCGCAGACCGAGTTGGGGCCA--CTTGATCACAGTCCGGGACC	849	
RESULT 2			
BI859740			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 884)			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
Contact: Robert Strausberg, Ph.D.			
Email: c9aps-f@mail.nih.gov			
Tissue Procurement: DCTD/DTF			
cDNA Library Preparation: Life Technologies, Inc.			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: L1AM1209 row: c column: 11			
High quality sequence stop: 865.			
FEATURES			
source			
1..884			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:539596"			
/tissue_type="mammary adenocarcinoma, cell line"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_87"			
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;			
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.			
Average insert size 1.383 kb. Library enriched for			
full-length clones and constructed by Life Technologies.			
Note: this is a NIH_MGC Library."			
ORIGIN			
Query Match 56.9%; Score 713.2; DB 4; Length 884;			
Best Local Similarity 93.0%; Pred. No. 4.4e-184;			
Matches 824; Conservative 0; Mismatches 28; Indels 34; Gaps 6;			
QY	289 GACGCTGTGTACCGGGGCTGCTGGAGCATCTGCGCAGCCAGAGCACAAGAAAGCAAC	348	
Db	1 GACGCTGTGTGTACCGGGGCTGCTGGAGCATCTGCGCAGCCAGAGCACAAGAAAGCAAC	60	
QY	349 CAACAAATTTCTGTGGGAGAAACAAAGCTGAGTCCAGATGAAGAGAGTTTCTTGGTCAC	408	
Db	61 CAACAAATTTCTGTGGGAGAAACAAAGCTGAGTCCAGATGAAGAGAGTTTCTTGGTCAC	120	
QY	409 TCCCGAGATTTATGCGGATTCAGAAATCCATGTAAGAGTTTGGATTCCTATGAAGA	468	
Db	121 TCCCGAGATTTATGCGGATTCAGAAATCCATGTAAGAGTTTGGATTCCTATGAAGA	180	
QY	469 AAAGAGGATTAAGTGAATCAAGAGATGCGCAGCTCAGATCCGTGAGGTGGAGCAGCGC	528	
Db	181 AAAGAGGATTAAGTGAATCAAGAGATGCGCAGCTCAGATCCGTGAGGTGGAGCAGCGC	240	
QY	529 ACAGAGGTGTTGCTGTCTTAGAG-----CCTC	560	
Db	241 ACAGAGGTGTTGCTGTCTTAGAGTTGGTTTCCCTCGAGGATTCAGACACCTC	300	
QY	561 AGGCACTGCCAGCCAGAGAGGCTCTTCAGACCTTAGAAGCTGGAAGGATGAACA	620	
Db	301 AGGCACTGCCAGCCAGAGAGGCTCTTCAGACCTTAGAAGCTGGAAGGATGAACA	360	
QY	621 GCCTAGTAGTTTCCAGCTTACAGCAGCCCTCAAATTTGGAACCTGCCACAGCTCCAGAGC	680	
Db	361 GCCTAGTAGTTTCCAGCTTACAGCAGCCCTCAAATTTGGAACCTGCCACAGCTCCAGAGC	420	
QY	681 TTGACTGGATGAGAGCAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCAGAG	740	
Db	421 TTGACTGGATGAGAGCAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCAGAG	480	
QY	741 TTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATATCATTTG	800	
Db	481 TTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATATCATTTG	540	
QY	801 CTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAA--GGAAACA	858	
Db	541 CTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAACGGAACA	600	

QY 859 GAAGTTGAATAAATCCCCCAGACCGAGTTGGGGCCAACTTTGATCAGCTCCAGGAC 918  
DB |||||  
601 GAAGTTGAATAAATCCCCCAGACCGAGTTGGGGCCAACTTTGATCAGCTCCAGGAC 660  
QY 919 CAGTGCAGGCTGGCTGCCCTCTTTTGGGGCCCGCTCTGGAATAATGGACG-CGGCTGGCAG 977  
DB |||||  
661 CAGTGCAGGCTGGCTGCCCTCTTTT-GGGCGGCTCTGGAATAATGGACGCGCGTTGGCAG 719  
QY 978 TCCAGACATCAATCTCAAAATCGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAA 1037  
DB |||||  
720 TCCAGACATCAA-ITCAAAATCGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAA 778  
QY 1038 AAGCTAATCATGCTCTCTACCAATACCATCAGGCTAAAGCCCAAGTCACCAACCC 1097  
DB |||||  
779 AAG-TAATCATGCTCTCTTACCAATACCATCAGGCTAAAGCCCAAGTCACCAACCCCT 837  
QY 1098 TATTATACCTTCCACCAAAATCTTTATCATCTGCTTTCTTAGGAA 1143  
DB |||||  
838 TATTATACCTTCCACCAAAATCTTTATCATCTGCTTTCTTAGGAA 883

RESULT 3  
BG403663  
LOCUS  
DEFINITION BG403663 1066 bp mRNA linear EST 12-MAR-2001  
mRNA sequence.  
ACCESSION BG403663.1 GI:13297111  
VERSION  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10433 row: j column: 21  
High quality sequence stop: 735.  
Location/Qualifiers  
1. .1066  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone IMAGE:4526324"  
/tissue type="transitional cell papilloma, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH MGC 93"  
/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dr primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

FEATURES  
source

RESULT 4  
BQ006160/c  
LOCUS  
DEFINITION BQ006160 714 bp mRNA linear EST 26-MAR-2002  
UT-H-E11-ayy-j-22-0-UI.s1 NCI CGAP\_E11 Homo sapiens CDNA clone  
IMAGE:5845365 3', mRNA sequence.  
ACCESSION BQ006160  
VERSION BQ006160.1 GI:19731060  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found

## ORIGIN

Query Match 55.4%; Score 694; DB 4; Length 1066;  
Best Local Similarity 97.3%; Pred. No. 8.6e-179;  
Matches 759; Conservative 0; Mismatches 15; Indels 6; Gaps 5;  
QY 99 GCAGCTTACAGCCGACAGCAGCCAGGAGCTGAAGAGGCTTTTGGAGGCTCTGCG 158  
DB |||||  
1 GCAGCTTACAGCCGACAGCA-CCAGCGCGAGCTGAAGAGGCTTTGGAGGCTCTGCG 59

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 The following repetitive elements were found in this cDNA  
 sequence: 1-41, >AT rich/Low\_complexity  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

## Location/Qualifiers

1..714  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5845365"  
 /tissue\_type="Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI\_CGAP\_E11"  
 /note="Organ: Left Pelvis; Vector: p7T73-Pac (Pharmacia)  
 with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP E11 is a normalized cDNA library containing the  
 following tissue(s): Chondrosarcoma. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into p7T73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 ACACCTGGCAC.  
 TAG\_TISSUE=chondrosarcoma  
 TAG\_LIB=UI-H-E11  
 TAG\_SEQ=ACACTGGCAC"

## ORIGIN

Query Match 49.2%; Score 616.8; DB 5; Length 714;  
 Best Local Similarity 98.1%; Pred. No. 1.2e-157;  
 Matches 677; Conservative 0; Mismatches 7; Indels 6; Gaps 5;

Qy 497 GCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGTTCGCTCTCTAGAG 556  
 Db 703 GCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGTTCGCTCTCTAGAG 644

Qy 557 CCTCAGCAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGAGCTGGAAAGGGATG 616  
 Db 643 CCTCAGCAGTGCAGACCCAGAGAGG-GGCTCTTCAGCACCTAGAGAGCTGGAAAGGGATG 585

Qy 617 AACAGCCAGTAGCTCCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 676  
 Db 584 AACAGCCAGTAGCTCCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 525

Qy 677 GAGCTTGACTGGATGGAGCAGACCATCTTCACATTCATATGGCCATCAGGATATACCA 736  
 Db 524 GAGCTTGACTGGATGGAGCAGACCATCTTCACATTCATATGGCCATCAGGATATACCA 465

Qy 737 GGAGTTGGTAAATCCACTCAGTGCACACCTCCCTGGATGATCCTCAAGATGAAGATAC 796  
 Db 464 GGAGTTGGTAAATCCACTCAGTGCACACCTCCCTGGATGATCCTCAAGATGAAGATAC 405

Qy 797 ATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAAAA 856  
 Db 404 ATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAAAA 345

Qy 857 CAGAGTTGAAAAAATCCCGCCAGACCGAGTGGGGCCAACTTGTATCACACTCCAGG 916  
 Db 344 CAGAGTTGAAAAAATCCCGCCAGACCGAGTGGGGCCAACTTGTATCACACTCCAGG 285

Qy 917 ACCAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGTCTTGGAAATTAATGGACGCCCTGGCA 976  
 Db 284 ACCAGTGCAGGCTGGCTGCCCTCTTTT-GGCGCGGCTCTGGAAATTAATGGACGCCCTGGCA 226

Qy 977 GTCCAGACATCAATCCAAAATCGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAA 1036  
 Db 225 GTCCAGACATCAA-TTCAAAATCGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAA 167

Qy 1037 AAAGCTAATCATGCTCTCTACCACTACCACTAGGCTAAAGCCAAAGTCAACCAACCC 1096  
 Db 166 AAAGCTAATCATGCTCTCTACCACTACCACTAGGCTAAAGC-AAAGTCAACCAACCC 109  
 Qy 1097 CTATTATACCTTCACCCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATACTCA 1156  
 Db 108 CTATTATACCTTCAC-CCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATACTCA 50  
 Qy 1157 TTCAATTTGATTTAATAAAGTTTATTATTTTC 1186  
 Db 49 TTCAATTTGATTTAATAAAGTTTATTATTTTC 20

## RESULT 5

BI769735

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM11512 row: h column: 09

High quality sequence stop: 790.

Location/Qualifiers

1..816

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5204408"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_122"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;

Site 1: NotI; Site 2: EcoRV (destroyed); RNA source

anonymous pool of 24 week female lung, 16 week female

spleen, and 20-22 week male spleens. Library is oligo-dT

primed and directionally cloned (EcoRV site is destroyed

upon cloning). Average insert size 1.4 kb. Insert size

range 1-3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 026. Note:

this is a NIH\_MGC Library."

## ORIGIN

Query Match 46.8%; Score 587; DB 4; Length 816;  
 Best Local Similarity 91.6%; Pred. No. 1.8e-149;  
 Matches 721; Conservative 0; Mismatches 5; Indels 61; Gaps 7;

Qy 81 CTTTCTTCTGTCGTCGC-GGCACGTTTACAGCCGCAAGCACCCAGCGGAGTGAAGGAG 139  
 Db 58 CTTTCTTCTGTCGTCGTCGCACGTTTACAGCCGCAAGCA-CCAGCGGAGTGAAGGAG 116  
 Qy 140 GCTTTTGAAGAGCTCTCTGCCCCAGGTGGAGCGGCCCGCAAGGCCATCCGCCCGCTCAG 199  
 Db 117 GCTTTGAGAGGCTCTCTGCCCCAGGTGGAGCGGCCCGCAAGGCCATCCGCCCGCTCAG 176



```
|||||
292 GCTCCAGGACGAGCTGGCTGCCCTCTTTT-GGCGCGCTCTGGAATATGAGCGC 234
QY CGCTGGCAGTCCAGACATCAACTCCAAAACAGCTGAGCTGCGACATCAAGAGCAGTCA 1028
Db CGCTGGCAGTCCAGACATCAA-FTCAAAACAGCTGAGCTGAGCAATGAAGAGCAGTCA 175
QY TACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAGTCAA 1088
Db TACAGAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC--AAAGTCA 117
QY CMAACCCCTATATACCTTCCACCACCAAAATCTTTATCATGCTCTTTCTTAGGAAACAGA 1148
Db ACAAAACCCCTATATACCTTCCA-CCAAATCTTTATCATGCTCTTTCTTAGGAAACAGA 58
QY CATACCTCATTCATTTGATTAATAAGTTTATTTTTC 1186
Db CATACCTCATTCATTTGATTAATAAGTTTATTTTTC 20
```

```
RESULT 7
BQ007036/c
LOCUS
DEFINITION
  BQ007036 758 bp mRNA linear EST 26-MAR-2002
  UI-H-EII-azb-p-24-0-UI.s1 NCI_CGAP_EII Homo sapiens cDNA clone
  IMAGE:5846663 3', mRNA sequence.
ACCESSION
  BQ007036
VERSION
  BQ007036.1 GI:19731936
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 758)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs@mail.nih.gov
  Tissue Procurement: Dr. Jose Mercuende
  CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Clone distribution information can be found
  through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
  The following repetitive elements were found in this cDNA
  sequence: 1-41, 2AT-rich/Low_complexity
  Seq primer: M13 FORWARD
  POLYA=Yes.
```

```
FEATURES
  source
  1..758
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5846663"
  /tissue_type="Chondrosarcoma"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies)"
  /clone_lib="NCI_CGAP_EII"
  /note="Organ: Left Pelvis; Vector: pTT3-Pac (Pharmacia)
  with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
  NCI_CGAP_EII is a normalized cDNA library containing the
  following tissue(s): Chondrosarcoma. The library was
  constructed according to Bonaldo, Lennon and Soares,
  Genome Research, 6:791-806, 1996. First strand cDNA
  synthesis was primed with an oligo-dT primer containing a
  Not I site. Double stranded cDNA was ligated to an EcoR I
  adaptor, digested with Not I, and cloned directionally
  into pTT3-Pac vector. The oligonucleotide used to prime
  the synthesis of first-strand cDNA contains a library tag
  sequence that is located between the Not I site and the
  (dT)18 tail. The sequence tag for this library is
  ACACTTGAC.
```

```

TAG_ISSUE=chondrosarcoma
TAG_LIB=UI-H-EII
TAG_SEQ=ACACTTGAC"
ORIGIN
  Query Match 46.0%; Score 576.4; DB 5; Length 758;
  Best Local Similarity 94.4%; Pred. No. 1.5e-146;
  Matches 706; Conservative 0; Mismatches 6; Indels 36; Gaps 9;
QY 466 AGAAAAGAGGATAAAGTGATCAAGGAGATGCGACGCTCAGATCCGTGAGGTGGAGCAGAG 525
Db 758 AGAAAAGGAGGATAAAGTGATCAAGGAGAT-GCAGCTCAGATCCGTGA-GTGGAGCAGAG 701
QY 526 CCGACAGAGGTTGGTTCGTCTCTTAGAG-----CC 558
Db 700 CCGACAGAGGT-GTTCGGTCTCTTAGAGGTGGTTCCTCGGAGGATCCAGACCACC 642
QY 559 TCAGGCAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAA 618
Db 641 TCAGGCAGTCCAGACCCAGAGAGG-GGCTCTTCAGCACCTAGAACCTGGAAGGGATGAA 583
QY 619 CAGCCAAAGTAGTCTTCAGGCTTACAGCAGCCCTCAAATTTGGACCTGCGACCGCTCCAGA 678
Db 582 CAGCCAAAGTAGTCTTCAGGCTTACAGCAGCCCTCAAATTTGGACCTGCGACCGCTCCAGA 523
QY 679 GCTTGACCTGGATGGAGAGCAGGACCATCTCTGACATTCATTTGGCCATCAGGATATACCAGG 738
Db 522 GCTTGACCTGGATGGAGAGCAGGACCATCTCTGACATTCATTTGGCCATCAGGATATACCAGG 463
QY 739 AGTTGGTAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGATACAT 798
Db 462 AGTTGGTAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGATACAT 403
QY 799 TGCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCTTAAAGAAAAAGGAAAAACA 858
Db 402 TGCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCTTAAAGAAAAAGGAAAAACA 343
QY 859 GAAAGTTGAAAAAACTCCGCCAGACCGAGTTGGGCCCAACTTTGATACACAGTCCAGAC 918
Db 342 GAAAGTTGAAAAAACTCCGCCAGACCGAGTTGGGCCCAACTTTGATACACAGTCCAGAC 283
QY 919 CAGTCAGGCTGGCTGGCTCTTTTGGCCCGCTCTGGAATAATGGACCGCGCTGGCAGT 978
Db 282 CAGTCAGGCTGGCTGGCTCTTTT-GGCCGCGCTCTGGAATAATGGACCGCGCTGGCAGT 224
QY 979 CCAGACATCAACTCCAAAACCTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAAA 1038
Db 223 CCAGACATCAA-TTCAAAACCTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAAA 165
QY 1039 AGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCCAAAGTCAACCAACCCCT 1098
Db 164 AGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGC--AAAGTCAACCAACCCCT 107
QY 1099 ATTATACCTTCCACCCCAAAATTTTATCATTTGCTTTCTTAGGAAACAGACATCTCATT 1158
Db 106 ATTATACCTTCCA-CCAAATTTCTTATCATTTGCTTTCTTAGGAAACAGACATCTCATT 48
QY 1159 CATTTGATTTAATAAGTTTATTTTTC 1186
Db 47 CATTTGATTTAATAAGTTTATTTTTC 20
RESULT 8
BQ682005/c
LOCUS
DEFINITION
  BQ682005 760 bp mRNA linear EST 27-FEB-2002
  UI-E-E01-aiw-c-24-0-UI.s1 UI-E-E01 Homo sapiens cDNA clone
  UI-E-E01-aiw-c-24-0-UI 3', mRNA sequence.
ACCESSION
  BQ682005
VERSION
  BQ682005.1 GI:18991901
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```









QY 440 ATGGTGAAGGTTTGGATTCTATGAAGAAGGAGGATAAGTCAAGAGATGGCA 499  
 |||||  
 Db 396 ATGGTGAAGGTTTGGATTCTATGAAGAAGGAGGATAAGTCAAGAGATGGCA 455  
 |||||  
 QY 500 GCTCAGATCCGTGAGTGGAGCAGACAGAGAGGTGGTTGGTCTGTCTTAGAGCT 559  
 |||||  
 Db 456 GCTCAGATCCGTGAGTGGAGCAGACAGAGAGGTGGTTGGTCTGTCTTAGAGCT 515  
 |||||  
 QY 560 CAGGAGTGCAGACCCAGAGAGAGGCTCTTTCAGCACTAGAACCTGGAAGGGATGAAC 619  
 |||||  
 Db 516 CAGGAGTGCAGACCCAGAGAGAGGCTCTTTCAGCACTAGAACCTGGAAGGGATGAAC 575  
 |||||  
 QY 620 AGCCAAGTAGCTTCCAGCTTACAGAGCCCTCAAA--TTTGGACCTGCCACAGCTCCAG 677  
 |||||  
 Db 576 AGCCAAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCAG 635  
 |||||  
 QY 678 AGCTT 682  
 |||||  
 Db 636 AGCTT 640

RESULT 11  
 BU608980/c  
 LOCUS  
 DEFINITION UI-CF-FNO-aer-m-01-0-UI.s1 743 bp mRNA linear EST 21-FEB-2003  
 UI-CF-FNO-aer-m-01-0-UI 3', mRNA sequence.  
 ACCESSION BU608980  
 VERSION  
 KEYWORDS  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 743)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: McCray, PB  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.reagen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 1-47, >AT rich#Low\_complexity  
 Seq primer: M13 FORWARD  
 POLYA-Yes.

FEATURES  
 source  
 1. 743  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-FNO-aer-m-01-0-UI"  
 /tissue\_type="Human Lung Epithelial cells"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-FNO"  
 /note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a  
 modified polylinker; Site1: EcoR I; Site2: Not I;  
 UI-CF-FNO is a subtracted cDNA library derived from two  
 normalized Human lung epithelial cell libraries (BN1 and  
 DU1) The library was subtracted according to according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:  
 bento-soares@uiowa.edu  
 TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS  
 6hr to LPS 24h  
 TAG LIB=UI-CF-FNO  
 TAG\_SEQ=CTGCTCAGGT"  
 ORIGIN  
 Query Match 44.6%; Score 558.4; DB 5; Length 743;  
 Best Local Similarity 98.1%; Pred. No. 1.3e-141;  
 Matches 618; Conservative 0; Mismatches 6; Indels 6; Gaps 5;  
 QY 557 CCTCAGGAGTGCAGACCCAGAGAGGCTCTTTCAGCACTAGAACCTGGAAGGGATG 616  
 |||||  
 Db 649 CCTCAGGAGTGCAGACCCAGAGAGGG-TCTTTCAGCACCTAGAACCTGGAAGGGATG 591  
 |||||  
 QY 617 AACAGCCAAAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 676  
 |||||  
 Db 590 AACAGCCAAAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCA 531  
 |||||  
 QY 677 GAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCA 736  
 |||||  
 Db 530 GAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCA 471  
 |||||  
 QY 737 GGAAGTTGATTAACATCCATCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATAC 796  
 |||||  
 Db 470 GGAAGTTGATTAACATCCATCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATAC 411  
 |||||  
 QY 797 ATTGCTGGGACCAAGAAATAGGACCATCTCTGACAGAAATTTCTTAAAGAAAGGAAAAA 856  
 |||||  
 Db 410 ATTGCTGGGACCAAGAAATAGGACCATCTCTGACAGAAATTTCTTAAAGAAAGGAAAAA 351  
 |||||  
 QY 857 CAGAAGTTGAAAAAATCTCCGCCACAGCGAGTTGGGGCCAACTTTTGATCAGAGCTCCAGG 916  
 |||||  
 Db 350 CAGAAGTTGAAAAAATCTCCGCCACAGCGAGTTGGGGCCAACTTTTGATCAGAGCTCCAGG 291  
 |||||  
 QY 917 ACCAGTCAAGCTGGCTGGCTCTTTTGGGCGCGTCTGGAAATATGGACGCGCTGGCA 976  
 |||||  
 Db 290 ACCAGTCAAGCTGGCTGGCTCTTTT-GGCGCGTCTGGAAATATGGACGCGCTGGCA 232  
 |||||  
 QY 977 GTCCAGACATCACTCCAAAATCAAGCTGAGCAATGAAGAAGCAGTCAATACAGAAA 1036  
 |||||  
 Db 231 GTCCAGACATCAA-TTCAAAAATCAAGCTGAGCAATGAAGAAGCAGTCAATACAGAAA 173  
 |||||  
 QY 1037 AAAGCTAATCATGCTCTCTACCAACTACCACTGAGGCTAAAGCCAAAGTCAACCAAGCCC 1096  
 |||||  
 Db 172 AAAGCTAATCATGCTCTCTACCAACTACCACTGAGGCTAAAGC--AAAGTCAACCAAGCCC 115  
 |||||  
 QY 1097 CTATTATACCTTCCACCCAAATCTTTTATCATTTCTTTCTTAGGAAACAGACATACTCA 1156  
 |||||  
 Db 114 CTATTATACCTTCCA-CCAAATCTTTTATCATTTCTTTCTTAGGAAACAGACATACTCA 56  
 |||||  
 QY 1157 TTCAATTTGATTTAATAAAGTTTATTTTTC 1186  
 |||||  
 Db 55 TTCAATTTGATTTAATAAAGTTTATTTTTC 26  
 |||||  
 RESULT 12  
 BU608895/c  
 LOCUS  
 DEFINITION UI-CF-FNO-aer-1-18-0-UI.s1 747 bp mRNA linear EST 21-FEB-2003  
 UI-CF-FNO-aer-1-18-0-UI 3', mRNA sequence.  
 ACCESSION BU608895  
 VERSION  
 KEYWORDS EST.  
 SOURCE BU608895.1 GI:23275110  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 747)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery

JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.reagen.com) or from Open Biosystems  
(www.openbiosystems.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-47, >AT rich#Low\_complexity  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source  
1. 747  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-FNO-aeg-1-18-0-UI"  
/tissue\_type="Human Lung Epithelial cells"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-FNO"  
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-CF-FNO is a subtracted cDNA library derived from two  
normalized Human lung epithelial cell libraries (EVI and  
DUI) The library was subtracted according to according to  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. For additional information, contact:  
bento-soares@uiowa.edu  
TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS  
6hr To LPS 24h  
TAG LIB=UI-CF-FNO  
TAG\_SEQ=CTGCTCAGGT"

ORIGIN  
Query Match 44.6%; Score 558.4; DB 5; Length 747;  
Best Local Similarity 98.1%; Pred. No. 1.3e-141;  
Matches 618; Conservative 0; Mismatches 6; Indels 6; Gaps 5;

QY 557 CCTCAGGCAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGGGATG 616  
|||||  
DB 649 CCTCAGGCAGTCCAGACCCAGAGAGGG-TCTTCAGCACCTAGAGCTGGAAGGGATG 591  
|||||  
QY 617 AACAGCCAGTAGTTCAGCTTACAGAGCCCTCAAAATTGGACTGCTGCCACCACTCCA 676  
|||||  
DB 590 AACAGCCAGTAGTTCAGCTTACAGAGCCCTCAAAATTGGACTGCTGCCACCACTCCA 531  
|||||  
QY 677 GAGCTTGACTGATGAGACAGACCATCTCTGACATTCATTGGCCATCAGGATATACCA 736  
|||||  
DB 530 GAGCTTGACTGATGAGACAGACCATCTCTGACATTCATTGGCCATCAGGATATACCA 471  
|||||  
QY 737 GGAGTTGGTAAATCACTCAGGTGCCACACCTCCCTGGATGATCAAGATGAAGAATAC 796  
|||||  
DB 470 GGAGTTGGTAAATCACTCAGGTGCCACACCTCCCTGGATGATCAAGATGAAGAATAC 411  
|||||  
QY 797 ATTGCTGGGAACCAAGAAATAGGACATCTCTATGAAGAAATTTCTTAAAGAAAAGAAAA 856  
|||||  
DB 410 ATTGCTGGGAACCAAGAAATAGGACATCTCTATGAAGAAATTTCTTAAAGAAAAGAAAA 351  
|||||  
QY 857 CAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGG 916  
|||||  
DB 350 CAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGG 291  
|||||  
QY 917 ACCAGTGCAGGCTGGCTGCCCTCTTTTGGGCCGCGTCTTGAATAATGGACGCCGCTGGCA 976  
|||||

Db 290 ACCAGTGCAGGCTGGCTGCCCTCTTTT-GGCGCGCTCTGGAATATGACGCCGCTGGCA 232  
|||||  
QY 977 GTCCAGACATCAACTCCAAACTGAAGCTGAGCAATGAAGAGCAGTCAACATACAGAAA 1036  
|||||  
Db 231 GTCCAGACATCAA-TTCAAACTGAAGCTGAGCAATGAAGAGCAGTCAACATACAGAAA 173  
|||||  
QY 1037 AAAGTAATCATGCTCTCTACCAACTACCATGAGGCTAAAGCCAAAGTCAACCAAAACC 1096  
|||||  
Db 172 AAAGTAATCATGCTCTCTACCAACTACCATGAGGCTAAAGC-AAAGTCAACAAACC 115  
|||||  
QY 1097 CTATTATACCTTCCACCCAAATCTTTATCATCTGTCTTTTAGGAAAACAGACATACCTCA 1156  
|||||  
Db 114 CTATTATACCTTCCA-CCAAATCTTTATCATCTGTCTTTCTTTAGGAAAACAGACATACCTCA 56  
|||||  
QY 1157 TTCAATTCATTTAATAAAGTTTATTTTTC 1186  
|||||  
Db 55 TTCAATTCATTTAATAAAGTTTATTTTTC 26  
|||||

RESULT 13  
BM684305/c  
LOCUS  
DEFINITION  
UI-E-EJ1-aj1-1-23-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone  
UI-E-EJ1-aj1-1-23-0-UI 3', mRNA sequence.  
ACCESSION  
BM684305  
VERSION  
BM684305.1 GI:18994201  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 656)  
AUTHORS  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
PUBMED  
8889548  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.reagen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-47, >AT rich#Low\_complexity  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
source  
1. 656  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ1-aj1-1-23-0-UI"  
/tissue\_type="fetal eyes, lens, eye anterior segment,  
optic nerve, retina, Retina Foveal and Macular, RPE and  
Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-E-EJ1"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-EJ1 is a subtracted cDNA library constructed  
according to Ronaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAAAGTG; retina, CCGCG; Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG\_L1B=UI-E-EJ1 TAG\_SEQ=GTCC"

ORIGIN		Query Match	43.8%; Score 548.6; DB 4; Length 656;
		Best Local Similarity	96.8%; Pred. No. 6e-139;
		Matches 613; Conservative 0; Mismatches 14; Indels 6; Gaps 5;	
Qy	554	GAGCTCAGGAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAACGCTGGAAGGG	613
Db	552	GAGAGCTCAGGATGCCAGACCCAGAGA-GGCTCTTCAGCACCTAGAACGCTGGAAGGG	594
Qy	614	ATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGGCCCTCAAAATTTGGACCTGCCACAGCT	673
Db	593	ATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGGCCCTCAAAATTTGGACCTGCCACAGCT	534
Qy	674	CCAGAGCTTGATGGAGAGAGGACCATCTCTGCATCTTTCGATTCATTTGGCCATCAGGATATA	733
Db	533	CCAGAGCTTGATGGAGAGAGGACCATCTCTGCATCTTTCGATTCATTTGGCCATCAGGATATA	474
Qy	734	CCAGAGTTGTATACATCCATCAGGTGCCACACTCCCTGGATGATCCAGATGAAGAA	793
Db	473	CCAGAGTTGTATACATCCATCAGGTGCCACACTCCCTGGATGATCCAGATGAAGAA	414
Qy	794	TACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAA	853
Db	413	TACATTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAAGAAAGGAA	354
Qy	854	AAACAGAGTTGAAAAAATCCCCAGAGCCGAGTGGGGCCAACTTTGATCAAGTCC	913
Db	353	AAACAGAGTTGAAAAAATCCCCAGAGCCGAGTGGGGCCAACTTTGATCAAGTCC	294
Qy	914	AGGACCATGAGGCTGGCTGCTCTTTTGGGGGGCTCTGGATTAATGACGCCGCTG	973
Db	293	AGGACCATGAGGCTGGCTGCTCTTTTGGGGGGCTCTGGATTAATGACGCCGCTG	235
Qy	974	GCAGTCCAGACATCAACTCCAAAATCTGAAGCTGCAGCAATCAAGAAGCAGTCAATACAG	1033
Db	234	GCAGTCCAGACATCAA-TTCAAACTGAGCTGCAGCAATCAAGAAGCAGTCAATACAG	176
Qy	1034	AAAAAGCTAATATGCTCTTACCAACTACCATGAGGCTAAAAGCCAAAGTCAACAAA	1093
Db	175	AAAAAGCTAATATGCTCTTACCAACTACCATGAGGCTAAAAGC--AAAGTCAACAAA	118
Qy	1094	CCCTATTATACCTTCCACCAATTTCTTTATCATTTCTTTCTTAGGAACACATAC	1153
Db	117	CCCTATTATACCTTCCCA-CCAAAATTTCTTTATCATTTCTTTCTTAGGAACACATAC	59
Qy	1154	TCATTTCATTTGATTTAATAAAGTTTATTTTTC	1186
Db	58	TCATTTCATTTGATTTAATAAAGTTTATTTTTC	26
RESULT 14			
BM929870			
LOCUS			
DEFINITION			
UI-E-EJ1-aj1-1-23-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone			
ACCESSION			
BM929870			
VERSION			
BM929870.1 GI:19389043			

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 596)
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
PUBMED	8889548
COMMENT	Contact: Soares, MB Coordinated Laboratory for Computational Genomics University of Iowa 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA Tel: 319 335 8250 Fax: 319 335 9565 Email: bento-soares@uiowa.edu Tissue Procurement: Dr. Gregg Hageman cDNA Library preparation: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
FEATURES	Seq primer: M13 REVERSE Location/Qualifiers 1..596 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="UI-E-EJ1-aj1-1-23-0-UI" /tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid" /dev_stage="fetal and adult" /lab_host="DH10B (Life Technologies) (T1 phage resistant)" /clone_lib="UI-E-EJ1" /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-E-EJ1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAAAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
ORIGIN	
Query Match	43.2%; Score 541.4; DB 5; Length 596;
Best Local Similarity	98.2%; Pred. No. 5.5e-137;
Matches 590; Conservative	0; Mismatches 6; Indels 5; Gaps 4;
QY	553 AGAGCCTCAGGAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAGAACGCTGGAAGG 612
Db	1 AGAGCCTCAGGAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTAGAACGCTGGAAGG 60
QY	613 GATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACGAG 672
Db	61 GATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACGAG 120
QY	673 TCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTTGGCATCAGGATAT 732
Db	121 TCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTTGGCATCAGGATAT 180

```
QY 733 ACCAGAGTTGTGTAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGA 792
Db |||||||
181 ACCAGAGTTGTGTAACATCCACTCAGGTGGCCACACCTCCCTGGATGATCCAAAGATGAAGA 240
QY 793 ATACATTTCTGGGAACCAAGAAATAGGACCATCCTATGAGNAATTTCTTAAAGAAAGGA 852
Db |||||||
241 ATACATTTCTGGGAACCAAGAAATAGGACCATCCTATGAGNAATTTCTTAAAGAAAGGA 300
QY 853 AAAACAGAGTTGAAAAAATCCCCCGAGACCGAGTTGGGGCCAACTTTGATCACAGTTC 912
Db |||||||
301 AAAACAGAGTTGAAAAAATCCCCCGAGACCGAGTTGGGGCCAACTTTGATCACAGTTC 360
QY 913 CAGGACCATGTCAGGCTGGCTGCCCTCTTTTGGGGCCGCTCTGGAATAATGGACCCCGCT 972
Db |||||||
361 CAGGACCATGTCAGGCTGGCTGCCCTCTTTTGGGGCCGCTCTGGAATAATGGACCCCGCT 419
QY 973 GCGAGTCCAGACATCAACTCCAAATCTGAAGCTGAGCAATGAAGAGCAAGTCAATACA 1032
Db |||||||
420 GCGAGTCCAGACATCAA-TTCAAACTGAAGCTGAGCAATGAAGAGCAAGTCAATACA 478
QY 1033 GAAAAAGCTAATCATGCTCTCTACCACTACCATGAGCTAAAAAGCCAAAGTCAACCAA 1092
Db |||||||
479 GAAAAAGCTAATCATGCTCTCTACCACTACCATGAGCTAAAAAGC--AAAGTCAACAA 536
QY 1093 ACCCTATTATACCTTCCACCAAAATCTTTATCATTTGTTCTTTAGGAAACAGACATA 1152
Db |||||||
537 ACCCTATTATACCTTCCA-CCAAATCTTTATCATTTGTTCTTTAGGAAACAGACATA 595
QY 1153 C 1153
Db 596 C 596
```

```
RESULT 15
BI046266/c
LOCUS
DEFINITION
MR3-FN0209-060201-016-f06 FN0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BI046266
VERSION
BI046266.1 GI:14452888
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 668)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
PUBMED
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-FN0209-
060201-016-f06&t3=2001-02-06&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 668.
Location/Qualifiers
1..668
/organism="Homo sapiens"
```

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="FN0209"
/note="Organ: prostate normal; Vector: puc18; Site: 1:
Smal; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
```

## ORIGIN

```
Query Match 42.4%; Score 531.8; DB 4; Length 668;
Best Local Similarity 91.6%; Pred. No. 2.4e-134;
Matches 513; Conservative 0; Mismatches 28; Indels 28; Gaps 4;

QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGGTGGT 540
Db |||||||
668 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGTGGT 609
QY 541 TCGGTCTGCTTAGAGCCT-----CAGGCAGTGCAGACCCAG 578
Db |||||||
608 TCGTCTGTTCTTAGAGCTTGGTTTCCCTCGGAGATCCAGACCAGCTGCCAGACCCAG 549
QY 579 AAGAGGGCTCTTCAGCACCTAGAAGCTGGAAGGGATGAACAGGCCAAGTAGCTTCCAGCT 638
Db |||||||
548 AAGAGGGCTCTTCAGCACCTAGAAGCTGGAAGGGATGAACAGGCCAAGTAGCTTCCAGCT 489
QY 639 TACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTGACCTGGATGGAGACAG 698
Db |||||||
488 TACAGAGCCCTCAAAATTTGGACCTGCCACAGCTCCAGAGCTTGACCTGGATGGAGACAG 429
QY 699 GACCATCTCGACATTTCAATTCGCCCATCAGGATATACACAGGAGTTGGTAAACATCCACTCAG 758
Db |||||||
428 GACCATCTCGACATTTCAATTCGCCCATCAGGATATACACAGGAGTTGGTAAACATCCACTCAG 369
QY 759 GTGCCACACCTCCCTCGGATGATCCAAGATGAAGAAATACATTGCTGGGAACCAAGAAATAG 818
Db |||||||
368 GTGCCACACCTCCCTCGGATGATCCAAGATGAAGAAATACATTGCTGGGAACCAAGAAATAG 309
QY 819 GACCATCTCGAAGAAATTTCTTAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCCCTCC 878
Db |||||||
308 GACCATCTCGAAGAAATTTCTTAAAGAAAAGGAAAAACAGAAAGTTGAAAAAATCCCTCC 249
QY 879 CAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTGCAGCTGGCTGGCTCC 938
Db |||||||
248 CAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTGCAGCTGGCTGGCTCC 189
QY 939 CTTTTGGGCCCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAGACATCAACTCCAAAAC 998
Db |||||||
188 CTTTT-GNCCGCGTCTGGAATAATGGACCGCGCTGGCAGTCCAGACATCAA-TTCAAAAAC 131
QY 999 TGAAGCTCGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGTCTCTTACC 1058
Db |||||||
130 TGAAGCTCGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGTCTCTTACC 71
QY 1059 AACTACCATGAGGCTAAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAAT 1118
Db |||||||
70 AACTACCATGAGGCTAAAAGCNA-----GTCACAAAACCCCTATTATACCTTCCACCTAGAC 15
QY 1119 TCCTTTATCA 1127
Db 14 TCCTTTCCA 6
```

Search completed: June 10, 2005, 20:27:43  
Job time : 2650.16 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:21:54 ; Search time 138.136 Seconds  
(without alignments)  
14842.245 Million cell updates/sec

Title: US-09-155-676B-4  
Perfect score: 1253  
Sequence: 1 cattggagtcacgcggtggc.....ggggggcgctaccacctttt 1253

Scoring table: IDENTITY NUC  
Gapop 10'0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278.4	22.2	816	4	US-09-949-016-3437
2	244.8	19.5	374	4	US-09-513-999C-22599
3	175.8	14.0	6623	4	US-09-949-016-16221
4	175.8	14.0	7031	4	US-09-949-016-15179
5	122.8	9.8	328	4	US-09-513-999C-34197
6	83.4	6.7	97	4	US-09-513-999C-15628
7	60.4	4.8	7218	1	US-08-232-463-14
8	46.8	3.7	2126	2	US-08-789-354-1
9	46.8	3.7	2126	3	US-09-110-937-1
10	46.8	3.7	2126	3	US-09-058-725B-1
11	46.8	3.7	2126	3	US-09-232-857-1
12	44.8	3.6	1236	2	US-08-933-115-1
13	44.8	3.6	1236	3	US-08-205-008-1
14	44.8	3.6	1236	4	US-09-206-115-1
15	43.8	3.5	1981	3	US-08-981-392-26
16	43.8	3.5	1981	4	US-09-908-322-26
17	43.2	3.4	1023	1	US-08-198-446B-1
18	43.2	3.4	1023	2	US-08-870-693-1
19	43.2	3.4	2150	1	US-08-198-446B-10
20	43.2	3.4	2150	2	US-08-870-693-10
21	43	3.4	852	1	US-08-075-533-21
22	43	3.4	852	2	US-08-948-176-21
23	43	3.4	852	5	PCT-US91-09160-21
24	42.4	3.4	637	4	US-09-387-286-25
25	41.8	3.3	470	4	US-10-101-464A-422
26	41.2	3.3	950	4	US-09-636-499-20
27	41	3.3	51259	3	US-08-781-891-209

c	28	41	3.3	51259	4	US-09-618-166-209	Sequence 209, App
	29	40.8	3.3	71278	4	US-09-949-016-11851	Sequence 11851, A
	30	40.8	3.3	71278	4	US-09-949-016-17563	Sequence 17563, A
c	31	40.8	3.3	168174	4	US-10-071-411A-63	Sequence 63, Appl
c	32	40.8	3.3	168273	4	US-10-071-411A-2	Sequence 2, Appl
	33	40.6	3.2	145	4	US-09-822-250A-6	Sequence 6, Appl
	34	40.6	3.2	145	4	US-10-034-350A-6	Sequence 6, Appl
	35	40.6	3.2	148	4	US-09-822-250A-7	Sequence 7, Appl
	36	40.6	3.2	148	4	US-10-034-350A-7	Sequence 7, Appl
	37	40.6	3.2	149	4	US-09-822-250A-8	Sequence 8, Appl
	38	40.6	3.2	149	4	US-10-034-350A-8	Sequence 8, Appl
	39	40.6	3.2	150	4	US-09-822-250A-9	Sequence 9, Appl
	40	40.6	3.2	150	4	US-10-034-350A-9	Sequence 9, Appl
	41	40.4	3.2	3751	4	US-09-140-378A-1	Sequence 1, Appl
	42	40.4	3.2	3803	4	US-09-023-655-1272	Sequence 1272, Ap
	43	40.4	3.2	3943	4	US-09-949-016-2495	Sequence 2495, Ap
	44	40.4	3.2	3943	4	US-09-949-016-2496	Sequence 2496, Ap
	45	40.4	3.2	3943	4	US-09-949-016-2497	Sequence 2497, Ap

ALIGNMENTS

RESULT 1  
US-09-949-016-3437  
; Sequence 3437, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE OF INVENTION: CLO01307  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3437  
; LENGTH: 816  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-3437

Query Match	22.2%	Score	278.4	DB	4	Length	816
Best Local Similarity	96.7%	Pred. No.	1.1e-71				
Matches	327	Conservative	0	Mismatches	6	Indels	5
		Gaps	4				
Qy	849	AGGAAAAAGAGAGTTGAAAAAATCCCCAGACCCAGTTGGGCGCACTTTGATCACA	908				
Db	1	AGGAAAAAGAGAGTTGAAAAAATCCCCAGACCCAGTTGGGCGCACTTTGATCACA	60				
Qy	909	GCTCCAGACAGTGACAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATAATGGAGCG	968				
Db	61	GCTCCAGACAGTGACAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATAATGGAGCG	119				
Qy	969	CGCTGGCAGTCCAGACATCAACTCCAAAATGAAAGCTGCAGCAATGAAGAACGAGTCACA	1028				
Db	120	CGCTGGCAGTCCAGACATCAATTTCAAAATGAAAGCTGCAGCAATGAAGAACGAGTCACA	178				
Qy	1029	TACGAAAAAGCTAATCATGCTCTTACCAATACCATGAGGCTAAAGCAAGTCAA	1088				
Db	179	TACGAAAAAGCTAATCATGCTCTTACCAATACCATGAGGCTAAAGCAAGTCAA	236				
Qy	1089	CCAAACCCCTATTATACCTTCCACCCAAATCTTTTATCATTTCTTTCTTAGGAAACAGA	1148				
Db	237	ACAAACCCCTATTATACCTTCCCA-CCAATCTTTATCATTTCTTTCTTAGGAAACAGA	295				
Qy	1149	CATACTCATTTGATTTAATAAAGTTTTATTTTC	1186				









ADDRESSEE: SmithKline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/058,725B  
FILING DATE: April 10, 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/789,354  
FILING DATE: 27-JAN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Han, William T.  
REGISTRATION NUMBER: 34,344  
REFERENCE/DOCKET NUMBER: GP50005-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5219  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2126 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-058-725B-1

Query Match 3.7%; Score 46.8; DB 3; Length 2126;  
Best Local Similarity 96.0%; Pred. No. 0.0049;  
Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1204 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCACTTTT 1253  
Db 53 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCAGTTT 4

RESULT 11  
US-09-232-857-1/c  
Sequence 1, Application US/09232857  
Patent No. 6159700  
GENERAL INFORMATION:  
APPLICANT: DOUGLAS, STEPHEN  
APPLICANT: WILLETTTE, ROBERT  
APPLICANT: AIYAR, NAMBI  
APPLICANT: ROMANIC, ANNE  
APPLICANT: KHANDOUDI, NASSIRAH  
APPLICANT: GOUT, BERNARD  
APPLICANT: AL-BARAZANJI, KAMAL  
APPLICANT: AMES, ROBERT S.  
APPLICANT: FOLEY, JAMES J.  
APPLICANT: SARAU, HENRY  
APPLICANT: CHAMBERS, JON K.  
APPLICANT: SHABON, USMAN  
APPLICANT: BERGSMAN, DEREK  
TITLE OF INVENTION: A METHOD OF FINDING AGONIST  
AND ANTAGONIST TO HUMAN AND RAT GPR14  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/232,857  
FILING DATE: 15-JAN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/789,354  
FILING DATE: 27-JAN-1997  
APPLICATION NUMBER: 60/074,075  
FILING DATE: 09-FEB-1998  
APPLICATION NUMBER: 09/058,725  
FILING DATE: 10-APR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F.  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GP-50005-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2126 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-232-857-1

Query Match 3.7%; Score 46.8; DB 3; Length 2126;  
Best Local Similarity 96.0%; Pred. No. 0.0049;  
Matches 48; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1204 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCACTTTT 1253  
Db 53 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGCGTACCCAGTTT 4

RESULT 12  
US-08-933-115-1/c  
Sequence 1, Application US/08933115  
Patent No. 5866369  
GENERAL INFORMATION:  
APPLICANT: Zalacain, Magdalena  
APPLICANT: Brown, James R.  
TITLE OF INVENTION: NOVEL AmpS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: US  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/933,115  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dickinson, Todd Q  
REGISTRATION NUMBER: 28,354  
REFERENCE/DOCKET NUMBER: GM10099

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-933-115-1
    Query Match          3.6%; Score 44.8; DB 2; Length 1236;
    Best Local Similarity 95.8%; Pred. No. 0.014; 2; Indels 0; Gaps 0;
    Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1203 AATCAAGCTTATCGATACCGTCTGACCTCGAGGGGGCGGTACCCCACT 1250
Db 48 AATCAAGCTTATCGATACCGTCTGACCTCGAGGGGGCGGTACCCCAAT 1

RESULT 13
US-09-205-008-1/c
; Sequence 1, Application US/09205008
; Patent No. 6297039
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; TITLE OF INVENTION: NOVEL Amps
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/205,008
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/933,115
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10099
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-205-008-1
    Query Match          3.6%; Score 44.8; DB 3; Length 1236;
    Best Local Similarity 95.8%; Pred. No. 0.014;
    Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1203 AATCAAGCTTATCGATACCGTCTGACCTCGAGGGGGCGGTACCCCACT 1250
Db 48 AATCAAGCTTATCGATACCGTCTGACCTCGAGGGGGCGGTACCCCAAT 1

RESULT 14
US-09-206-115-1/c
; Sequence 1, Application US/09206115
; Patent No. 6558919
; GENERAL INFORMATION:
; APPLICANT: Zalacain, Magdalena
; APPLICANT: Brown, James R.
; TITLE OF INVENTION: NOVEL Amps
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/206,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/933,115
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd O
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: GM10099
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-09-206-115-1
    Query Match          3.6%; Score 44.8; DB 4; Length 1236;
    Best Local Similarity 95.8%; Pred. No. 0.014;
    Matches 46; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1203 AATCAAGCTTATCGATACCGTCTGACCTCGAGGGGGCGGTACCCCACT 1250
Db 48 AATCAAGCTTATCGATACCGTCTGACCTCGAGGGGGCGGTACCCCAAT 1

RESULT 15
US-08-981-392-26/c
; Sequence 26, Application US/08981392
; Patent No. 6262025
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos Manuel Pinto
; APPLICANT: Lewis, Julian Hart
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Gray, Grace
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
; TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
```

```
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,392
; FILING DATE: 22-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Antler, Adriane M.
; REGISTRATION NUMBER: 32,605
; REFERENCE/DOCKET NUMBER: 7326-038
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1981 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-981-392-26

Query Match      3.5%; Score 43.8; DB 3; Length 1981;
Best Local Similarity 95.7%; Pred. No. 0.036;
Matches 45; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1204 ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCACT 1250
         |||||||||||||||||||||||||||||||||||||||||||
Db      48  ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGCGGTACCCCAAT 2
```

Search completed: June 10, 2005, 20:46:47  
Job time : 140.386 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 476.179 Seconds  
(without alignments)  
16311.995 Million cell updates/sec

Title: US-09-155-676B-4  
Perfect score: 1253  
Sequence: 1 cattggagtcacgctggc.....ggggggcgctaccactttt 1253

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 309930249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
- 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
- 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
- 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
- 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	451.8	36.1	3948	9	US-09-915-543-16
2	451.8	36.1	3948	15	US-10-322-579-16
3	451.8	36.1	3948	17	US-10-664-859-16
4	444.4	35.5	500	18	US-10-242-535A-45728
5	444.4	35.5	500	18	US-10-085-783A-45728
6	400.8	32.0	2556	10	US-09-746-783-205
7	256.4	20.5	352	9	US-09-796-692-3344

Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 45728, A  
Sequence 45728, A  
Sequence 205, App  
Sequence 3344, Ap

8	256.4	20.5	352	14	US-10-040-862-3344	Sequence 3344, Ap
9	256.4	20.5	352	17	US-10-057-475B-3344	Sequence 3344, Ap
10	256.4	20.5	352	17	US-10-154-884B-3344	Sequence 3344, Ap
11	256.4	20.5	352	19	US-10-764-324-3344	Sequence 3344, Ap
12	219.6	17.5	970	13	US-10-027-632-120356	Sequence 120356,
13	219.6	17.5	970	17	US-10-027-632-120356	Sequence 120356,
c 14	205	16.4	567	21	US-10-956-157-3788	Sequence 3788, Ap
15	205	16.4	567	21	US-10-956-157-3788	Sequence 3788, Ap
16	186.4	14.9	254	17	US-10-342-535A-52797	Sequence 52797, A
17	186.4	14.9	254	18	US-10-085-783A-52797	Sequence 52797, A
18	104.8	8.4	258	17	US-10-131-827-8756	Sequence 8756, Ap
19	45.8	3.7	864	14	US-10-198-846-3416	Sequence 3416, Ap
20	45	3.6	647	21	US-10-643-775-1354	Sequence 1354, Ap
21	44.4	3.5	437	20	US-10-357-930-31854	Sequence 31854, A
c 22	44.2	3.5	1488	19	US-10-437-963-42386	Sequence 42386, A
c 23	43.8	3.5	1981	9	US-09-908-322-26	Sequence 26, Appl
c 24	43.8	3.5	1981	10	US-09-783-931-26	Sequence 26, Appl
c 25	43.8	3.5	1981	20	US-10-877-563-16	Sequence 16, Appl
26	43.4	3.5	1896	19	US-10-437-963-90219	Sequence 90219, A
27	43.2	3.4	468	20	US-10-357-930-31864	Sequence 31864, A
28	43.2	3.4	468	20	US-10-357-930-40827	Sequence 40827, A
29	43.2	3.4	478	20	US-10-357-930-31862	Sequence 31862, A
30	43.2	3.4	495	20	US-10-357-930-31845	Sequence 31845, A
31	43.2	3.4	495	20	US-10-357-930-40811	Sequence 40811, A
c 32	43.2	3.4	575	14	US-10-198-846-4142	Sequence 4142, Ap
c 33	43.2	3.4	805	14	US-10-643-775-325	Sequence 7497, Ap
34	43	3.4	810	21	US-10-643-775-668	Sequence 668, App
35	43	3.4	810	21	US-10-300-341-25	Sequence 25, Appl
c 36	42.4	3.4	637	16	US-10-198-846-4856	Sequence 4856, Ap
37	42.2	3.4	860	14	US-10-101-464A-422	Sequence 422, App
38	41.8	3.3	470	14	US-10-864-252-422	Sequence 100273,
39	41.8	3.3	470	21	US-10-437-963-100273	Sequence 554, App
40	41.6	3.3	1664	19	US-10-643-775-554	Sequence 1030, Ap
41	41.4	3.3	526	21	US-10-643-775-554	Sequence 9085, Ap
42	41.2	3.3	783	21	US-10-643-775-1030	Sequence 20, Appl
43	41.2	3.3	824	14	US-10-198-846-9085	
c 44	41.2	3.3	950	15	US-10-229-358-20	
c 45	41	3.3	416	9	US-09-924-035A-641	

ALIGNMENTS

RESULT 1

US-09-915-543-16  
; Sequence 16, Application US/0915543  
; Publication No. US20020086986A1  
; GENERAL INFORMATION:  
; APPLICANT: BASLER, Konrad  
; APPLICANT: BRUNNER, Erich  
; APPLICANT: FROESCH, Barbara  
; APPLICANT: KRAMPS, Thomas  
; APPLICANT: PETER, Oliver  
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WIRELESS SIGNALING PATHWAY  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON  
; FILE REFERENCE: Q60361  
; CURRENT APPLICATION NUMBER: US/09/915,543  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: 60/221,502  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 3948  
; TYPE: DNA  
; ORGANISM: Human lgs-1  
; US-09-915-543-16

Query Match 36.1%; Score 451.8; DB 9; Length 3948;

Best Local Similarity 99.2%; Pred No. 6.8e-126;

Matches 475; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 81 CTTTCTTGTGTGTCG-GGCACGTTTACAGCCGACCCAGCCGACCTGAAGGAG 139

Db	3199	GTGGAGCGTATGTGCCGAAACACGAGCGATGCTGCTGTGCTGTGCTGCGCTGTGAG	32528
Qy	260	GTGCGGAAACACTGAGCCATGGAAACCTGCAGCGTGCTCTACGGGGGGTGCTGGAGCAT	319
Db	3259	GTGCGGAAACACTGAGCCATGGAAACCTGCAGCGTGCTCTACGGGGGGTGCTGGAGCAT	3318
Qy	320	CTGGCAGCCGACGACACAAGAACCAACAAATTCGTGGGAGACAACAAAGCTGAG	379
Db	3319	CTGGCAGCCGACGACACAAGAACCAACAAATTCGTGGGAGACAACAAAGCTGAG	3378
Qy	380	GTCCAGATGAACAGAGAAGTTTCTGTGTCACCTCCCCAGGATTATCGCGATTCAAGAAATCC	439
Db	3379	GTCCAGATGAACAGAGAAGTTTCTGTGTCACCTCCCCAGGATTATCGCGATTCAAGAAATCC	3438
Qy	440	ATGGTGAAGGTTTGATTCTCTATGAAGAAACGAGGATAAGTGATCAAGGAGATGGCA	499
Db	3439	ATGGTGAAGGTTTGATTCTCTATGAAGAAACGAGGATAAGTGATCAAGGAGATGGCA	3498
Qy	500	GCTCAGATCCGTGAGGTGGAGACGACCGACAGGAGTGGTTCGGTGCTGTCTTAGAGCC	558
Db	2499	CTCTGATCTCTGATGCTGAGTGAACCCGACAGGAGTGGTTCGGTGCTGTCTTAGAGAC	3557

RESULT 3  
US-10-664-859-16  
; Sequence 16, Application US/10664859  
; Publication No. US2004038901A1  
; GENERAL INFORMATION:  
; APPLICANT: BASLER, Konrad  
; APPLICANT: BRUNNER, Erich  
; APPLICANT: FROESCH, Barbara  
; APPLICANT: KRAMPS, Thomas  
; APPLICANT: PETER, Oliver  
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON  
; FILE REFERENCE: Q60361  
; CURRENT APPLICATION NUMBER: US/10/664,859  
; CURRENT FILING DATE: 2003-09-22  
; PRIOR APPLICATION NUMBER: US/09/915,543  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: 60/221,502  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 3948  
; TYPE: DNA  
; ORGANISM: Human lgs-1  
US-10-664-859-16

Db 3080 CCTTCTTTGTGGTCGCGGCACGTTTACAGCCGAAGCA-CCAGCGGACGCTGAAGGAG 3138  
QY 140 GCTTTTGAGAGGCTCCTGCCCCAGGTGGAGGGCGCCGCAAGGCCATCCGCGCGCTCAG 199  
Db 3139 GCTTTGGAGAGGCTCCTGCCCCAGGTGGAGGGCGCCGCAAGGCCATCCGCGCGCTCAG 3198  
QY 200 GTGGAGCGCTATGTGCCCGAACAAGAGCGATGCTGTGGTGTGCTGTGCTGCGGCTGTGAG 259  
Db 3199 GTGGAGCGCTATGTGCCCGAACAAGAGCGATGCTGTGGTGTGCTGTGCTGCGGCTGTGAG 3258  
QY 260 GTGCGGGAACACTGTGAGCCCATGAAACCCTGACGGTGTGTACGGGGGGCTGTGCGAGCAT 319  
Db 3259 GTGCGGGAACACTGTGAGCCCATGAAACCCTGACGGTGTGTACGGGGGGCTGTGCGAGCAT 3318  
QY 320 CTGGCCAGCCCCAGAGCACAGAAGCAACAAATAATTCTGTGGGGAGAACAAAGCTGAG 379  
Db 3319 CTGGCCAGCCCCAGAGCACAGAAGCAACAAATAATTCTGTGGGGAGAACAAAGCTGAG 3378  
QY 380 GTCCAGATGAAGAGAGATTCTGTGTCACCTCCCAGGATTATGCGCGATTCAAGAAAATCC 439  
Db 3379 GTCCAGATGAAGAGAGATTCTGTGTCACCTCCCAGGATTATGCGCGATTCAAGAAAATCC 3438  
QY 440 ATGGTGAAGGTTTGGATTCTTATGAGAAAAAGGAGGATAAAGTGATCAAGAGAGATGGCA 499  
Db 3439 ATGGTGAAGGTTTGGATTCTTATGAGAAAAAGGAGGATAAAGTGATCAAGAGAGATGGCA 3498  
QY 500 GCTCAGATCCGTGAGGTGGAGCAGCCGACAGGAGGTGCTCGGTCTGCTTAGAGCC 558  
Db 3499 GCTCAGATCCGTGAGGTGGAGCAGCCGACAGGAGGTGCTCGGTCTGCTTAGAGAC 3557

RESULT 2  
US-10-322-579-16  
; Sequence 16, Application US/10322579  
; Publication No. US20030114413A1  
; GENERAL INFORMATION:  
; APPLICANT: BASLER, Konrad  
; APPLICANT: BRUNNER, Erich  
; APPLICANT: FROESCH, Barbara  
; APPLICANT: KRAMPS, Thomas  
; APPLICANT: PETER, Oliver  
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON  
; FILE REFERENCE: Q60361  
; CURRENT APPLICATION NUMBER: US/10/322,579  
; CURRENT FILING DATE: 2002-12-19  
; PRIOR APPLICATION NUMBER: US/09/915,543  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: 60/221,502  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 3948  
; TYPE: DNA  
; ORGANISM: Human Igs-1  
; JS-10-322-579-16

Qy 380 GTCAGATGAAGAGAAAGTTCTGGTCACTCCCGAGGATTAATGCGCGATTCAAGAAATCC 439  
Db |||||  
Qy 3379 GTCCAGATGAAGAGAAAGTTCTGGTCACTCCCGAGGATTAATGCGCGATTCAAGAAATCC 3438  
Db |||||  
Qy 440 ATGGTGAAGGTTTGGATTCTCTATCAAGAAAGAGGATTAAGTCAATCAAGGATGCGCA 499  
Db |||||  
Qy 3439 ATGGTGAAGGTTTGGATTCTCTATCAAGAAAGAGGATTAAGTCAATCAAGGATGCGCA 3498  
Db |||||  
Qy 500 GCTCAGATCCGTGAGGTGAGCAGAGCCGACAGAGGAGTGGTTCGGTCTGTCTTAGAGCC 558  
Db |||||  
Qy 3499 GCTCAGATCCGTGAGGTGAGCAGAGCCGACAGAGGAGTGGTTCGGTCTGTCTTAGAGAC 3557  
Db |||||

## RESULT 4

US-10-242-535A-45728  
; Sequence 45728, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 45728  
; LENGTH: 500  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-45728

Query Match 35.5%; Score 444.4; DB 17; Length 500;  
Best Local Similarity 97.8%; Pred. No. 3.7e-124; Indels 5; Gaps 4;  
Matches 493; Conservative 0; Mismatches 6;  
Qy 683 GACTGGATGGAGACAGAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCGAGGTT 742  
Db |||||  
Qy 743 GGTAACTCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATTGCT 802  
Db |||||  
Qy 61 GGTAACTCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATTGCT 120  
Db |||||  
Qy 803 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 862  
Db |||||  
Qy 121 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 180  
Db |||||  
Qy 863 TTGAAAAAATCTCCCGCCAGACCGAGTTGGGCGCCAACTTTTGATCAAGTGAAGATACATTGCT 922  
Db |||||  
Qy 181 TTGAAAAAATCTCCCGCCAGACCGAGTTGGGCGCCAACTTTTGATCAAGTGAAGATACATTGCT 240  
Db |||||  
Qy 923 GCAGGCTGGTGCCTCTTTTGGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 982  
Db |||||  
Qy 241 GCAGGCTGGTGCCTCTTTTGGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 299  
Db |||||  
Qy 983 ACATCAACTCCAAACTGAAGCTGCAAGATGAAGAGCAGTCAATACAGAAAAAGCT 1042  
Db |||||  
Qy 300 ACATCAACTCCAAACTGAAGCTGCAAGATGAAGAGCAGTCAATACAGAAAAAGCT 358  
Db |||||  
Qy 1043 AATCATGCTCTCTACCACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 1102  
Db |||||  
Qy 359 AATCATGCTCTCTACCACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 416  
Db |||||

Qy 1103 TACCTTCCACCCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATACTATTCAATT 1162  
Db |||||  
Qy 417 TACCTTCCA-CCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATACTATTCAATT 475  
Db |||||  
Qy 1163 TGATTTAATAAAGTTTATTTTTC 1186  
Db |||||  
Qy 476 TGATTTAATAAAGTTTATTTTTC 499  
Db |||||

## RESULT 5

US-10-085-783A-45728  
; Sequence 45728, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 45728  
; LENGTH: 500  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-45728

Query Match 35.5%; Score 444.4; DB 18; Length 500;  
Best Local Similarity 97.8%; Pred. No. 3.7e-124; Indels 5; Gaps 4;  
Matches 493; Conservative 0; Mismatches 6;  
Qy 683 GACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCGAGGTT 742  
Db |||||  
Qy 1 GACTGGATGGAGACAGGACCATCTCTGACATTCATTGGCCATCAGGATATACCGAGGTT 60  
Db |||||  
Qy 743 GGTAACTCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATTGCT 802  
Db |||||  
Qy 61 GGTAACTCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATTGCT 120  
Db |||||  
Qy 803 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 862  
Db |||||  
Qy 121 GGGAAACCAAGAAATAGGACCATCTCTATGAAGAAATTTCTTAAAGAAAGGAAAAACAGAG 180  
Db |||||  
Qy 863 TTGAAAAAATCTCCCGCCAGACCGAGTTGGGCGCCAACTTTTGATCAAGTGAAGATACATTGCT 922  
Db |||||  
Qy 181 TTGAAAAAATCTCCCGCCAGACCGAGTTGGGCGCCAACTTTTGATCAAGTGAAGATACATTGCT 240  
Db |||||  
Qy 923 GCAGGCTGGTGCCTCTTTTGGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 982  
Db |||||  
Qy 241 GCAGGCTGGTGCCTCTTTTGGGCGCGCTCTGGAATAATGGACCGCGCTGGCAGTCCAG 299  
Db |||||  
Qy 983 ACATCAACTCCAAACTGAAGCTGCAAGATGAAGAGCAGTCAATACAGAAAAAGCT 1042  
Db |||||  
Qy 300 ACATCAACTCCAAACTGAAGCTGCAAGATGAAGAGCAGTCAATACAGAAAAAGCT 358  
Db |||||  
Qy 1043 AATCATGCTCTCTACCACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 1102  
Db |||||  
Qy 359 AATCATGCTCTCTACCACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTA 416  
Db |||||  
Qy 1103 TACCTTCCACCCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATACTATTCAATT 1162  
Db |||||  
Qy 417 TACCTTCCA-CCAAATTTCTTTATCATTTGCTTTCTTAGGAAACAGACATACTATTCAATT 475  
Db |||||  
Qy 1163 TGATTTAATAAAGTTTATTTTTC 1186  
Db |||||  
Qy 476 TGATTTAATAAAGTTTATTTTTC 499  
Db |||||

RESULT 6

US-09-746-783-205  
; Sequence 205, Application US/09746783  
; Publication No. US2003004935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESS: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 205:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2556 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 205:  
US-09-746-783-205

Query Match 32.0%; Score 400.8; DB 10; Length 2556;  
Best Local Similarity 97.4%; Pred. No. 1.8e-110;  
Matches 450; Conservative 0; Mismatches 7; Indels 5; Gaps 4;

QY	725	CAGGATATACAGGAGTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA	784
Db	2030	CAGGATATACAGGAGTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA	2089
QY	785	GATGAAGAATACATTCTCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCTTAAA	844
Db	2090	GATGAAGAATACATTCTCTGGGAACCAAGAAATAGGACCATCTATGAAGAATTTCTTAAA	2149
QY	845	GAAGAGGAAACACAGACTTGAAACATCCCTCCAGACCGAGTTGGGGCCAACTTGTAT	904
Db	2150	GAAGAGGAAACACAGAGTTGAAACATCCCTCCAGACCGAGTTGGGGCCAACTTGTAT	2209
QY	905	CACAGCTCCAGGACCAAGTGGCTGGCTCCCTCTTTTGGCCCGCTCTGGAATAATGG	964
Db	2210	CACAGCTCCAGGACCAAGTGGCTGGCTCCCTCTTTTGGCCCGCTCTGGAATAATGG	2268
QY	965	ACGCCGTGGCGAGTCAGACATCACTCCAAAACCTGAAGCTGCAGCAATGAAGACGAGT	1024

Db	2269	ACGCCGTGGCGAGTCAGACATCAA-TTCAAAACTGAAGCTGCAGCAATGAAGACGAGT	2327
QY	1025	CACATACAGAAAAAGCTAATCATCTCTCTACCACTACCATGAGGCTAAAAGCCAAAG	1084
Db	2328	CACATACAGAAAAAGCTAATCATCTCTCTACCACTACCATGAGGCTAAAAGC-AAA	2385
QY	1085	TCAACCAAAACCCCTATTATACCTTCCACCCTAAATCTTTATCATTTGTCTTTCTTAGGAAA	1144
Db	2386	GTCAACAAACCCCTATTATACCTTCCA-CCAAATCTTTATCATTTGTCTTTCTTAGGAAA	2444
QY	1145	CAGACATCTCATTCATTGATTTAATAAAGTTTTATTTTC	1186
Db	2445	CAGACATCTCATTCATTGATTTAATAAAGTTTTATTTTC	2486

RESULT 7

US-09-796-692-3344  
; Sequence 3344, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692

CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/222,903  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: 60/223,416  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/223,378  
PRIOR FILING DATE: 2000-08-07  
NUMBER OF SEQ ID NOS: 9597  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3344  
LENGTH: 352  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (92)  
OTHER INFORMATION: n=A, T, C or G  
NAME/KEY: unsure  
LOCATION: (187)  
OTHER INFORMATION: n=A, T, C or G  
NAME/KEY: unsure  
LOCATION: (235)  
OTHER INFORMATION: n=A, T, C or G  
NAME/KEY: unsure  
LOCATION: (242)  
OTHER INFORMATION: n=A, T, C or G  
NAME/KEY: unsure  
LOCATION: (318)



OTHER INFORMATION: n=A,T,C or G  
NAME/KEY: unsure  
LOCATION: (329)  
OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-3344

Query Match 20.5%; Score 256.4; DB 9; Length 352;  
Best Local Similarity 97.2%; Pred. No. 5.1e-67;  
Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;  
  
Qy 324 CCAGCCAGAGCAAGAAAGCAACCAACCAATTCGTGGGAGAAACAAAGCTGAGGTCC 383  
Db 1 CCAGCCAGAGCAAGAAAGCAACCAACCAATTCGTGGGAGAAACAAAGCTGAGGTCC 60  
  
Qy 384 AGATGAAGAGAGAGTTTCGTCTACTCCCGAGGATTTATCGCGATTCAAGAAATCCCATGG 443  
Db 61 AGATGAAGAGAGAGTTTCGTCTACTCCCGAGGATTTATCGCGATTCAAGAAATCCCATGG 120  
  
Qy 444 TGAAGGTTTGATTCCTATGAAGAAAGGAGGATTAAGTGATCAAGAGATGGCAGCTC 503  
Db 121 TGAAGGTTTGATTCCTATGAAGAAAGGAGGATTAAGTGATCAAGAGATGGCAGCTC 180  
  
Qy 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGAGGTGTTTCGGTCTGTCTTAGA-GCCTC 560  
Db 181 ACATCCNGTGGTGGAGCAGAGCCGACAGAGGTGTTTCGGTCTGTCTTAGANGCCTC 240  
  
Qy 561 AGGCAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGC-TGGAAGGGATGAAC 619  
Db 241 ANGCAAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGCTTGGAAGGGATGAAC 300  
  
Qy 620 AG-CCAAAGTAGCTTCCAGCTT 639  
Db 301 AGCCCAAGTAGCTTCCANCCTT 321

RESULT 8  
US-10-040-862-3344  
Sequence 3344, Application US/10040862  
Publication No. US20030078396A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
TITLE OF INVENTION: Hematological Malignancies  
FILE REFERENCE: 014058-013520US  
CURRENT APPLICATION NUMBER: US/10/040,862  
CURRENT FILING DATE: 2001-11-06  
PRIORITY APPLICATION NUMBER: US 60/186,126  
PRIORITY FILING DATE: 2000-03-01  
PRIORITY APPLICATION NUMBER: US 60/190,479  
PRIORITY FILING DATE: 2000-03-17  
PRIORITY APPLICATION NUMBER: US 60/200,545  
PRIORITY FILING DATE: 2000-04-27  
PRIORITY APPLICATION NUMBER: US 60/200,303  
PRIORITY FILING DATE: 2000-04-28  
PRIORITY APPLICATION NUMBER: US 60/200,779  
PRIORITY FILING DATE: 2000-04-28  
PRIORITY APPLICATION NUMBER: US 60/200,999  
PRIORITY FILING DATE: 2000-05-01  
PRIORITY APPLICATION NUMBER: US 60/202,084  
PRIORITY FILING DATE: 2000-05-04  
PRIORITY APPLICATION NUMBER: US 60/206,201  
PRIORITY FILING DATE: 2000-05-22  
PRIORITY APPLICATION NUMBER: US 60/218,950  
PRIORITY FILING DATE: 2000-07-14  
PRIORITY APPLICATION NUMBER: US 60/222,903  
PRIORITY FILING DATE: 2000-08-03  
PRIORITY APPLICATION NUMBER: US 60/223,416  
PRIORITY FILING DATE: 2000-08-04  
PRIORITY APPLICATION NUMBER: US 60/223,378

PRIOR FILING DATE: 2000-08-07  
PRIOR APPLICATION NUMBER: US 09/796,692  
PRIOR FILING DATE: 2001-03-01  
NUMBER OF SEQ ID NOS: 10467  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3344  
LENGTH: 352  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (92)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (187)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (235)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (242)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (318)  
OTHER INFORMATION: n=A,T,C or G  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (329)  
OTHER INFORMATION: n=A,T,C or G  
US-10-040-862-3344

Query Match 20.5%; Score 256.4; DB 14; Length 352;  
Best Local Similarity 97.2%; Pred. No. 5.1e-67;  
Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;  
  
Qy 324 CCAGCCAGAGCAAGAAAGCAACCAACCAATTCGTGGGAGAAACAAAGCTGAGGTCC 383  
Db 1 CCAGCCAGAGCAAGAAAGCAACCAACCAATTCGTGGGAGAAACAAAGCTGAGGTCC 60  
  
Qy 384 AGATGAAGAGAGAGTTTCGTCTACTCCCGAGGATTTATCGCGATTCAAGAAATCCCATGG 443  
Db 61 AGATGAAGAGAGAGTTTCGTCTACTCCCGAGGATTTATCGCGATTCAAGAAATCCCATGG 120  
  
Qy 444 TGAAGGTTTGATTCCTATGAAGAAAGGAGGATTAAGTGATCAAGAGATGGCAGCTC 503  
Db 121 TGAAGGTTTGATTCCTATGAAGAAAGGAGGATTAAGTGATCAAGAGATGGCAGCTC 180  
  
Qy 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGAGGTGTTTCGGTCTGTCTTAGA-GCCTC 560  
Db 181 ACATCCNGTGGTGGAGCAGAGCCGACAGAGGTGTTTCGGTCTGTCTTAGANGCCTC 240  
  
Qy 561 AGGCAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGC-TGGAAGGGATGAAC 619  
Db 241 ANGCAAGTCCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGCTTGGAAGGGATGAAC 300  
  
Qy 620 AG-CCAAAGTAGCTTCCAGCTT 639  
Db 301 AGCCCAAGTAGCTTCCANCCTT 321

RESULT 9  
US-10-057-475B-3344  
Sequence 3344, Application US/10057475B  
Publication No. US20040002068A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Clapper, Jonathan David

APPLICANT: Wang, Aijun  
APPLICANT: Ordonez, Nadia  
APPLICANT: Carter, Lauren  
APPLICANT: McNeill, Patricia Dianne  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
FILE REFERENCE: 014058-014402US  
CURRENT APPLICATION NUMBER: US/10/057,475B  
CURRENT FILING DATE: 2002-01-22  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: US 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: US 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: US 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 60/222,903  
PRIOR FILING DATE: 2000-08-03  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 10979  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3344  
LENGTH: 352  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(352)  
OTHER INFORMATION: n = g, a, c or t  
US-10-057-475B-3344

Query Match 20.5%; Score 256.4; DB 17; Length 352;  
Best Local Similarity 97.2%; Pred. No. 5.1e-67;  
Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

QY 324 CCAGCCGAGCAGCAGCAAGCAACCAAAATTCGTGGGAGAGCAAAAGCTGAGGTCC 383  
Db 1 CCAGCCGAGCAGCAGCAAGCAACCAAAATTCGTGGGAGAGCAAAAGCTGAGGTCC 60  
QY 384 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 443  
Db 61 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 120  
QY 444 TGAAGGTTTGGATTCTTATGAAGAAAAGGAGGATAAAGTGATCAAGGAGATGGCAGCTC 503  
Db 121 TGAAGGTTTGGATTCTTATGAAGAAAAGGAGGATAAAGTGATCAAGGAGATGGCAGCTC 180  
QY 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGGAGGTGGTTCGGTCTGTCTTAGA-GCCTC 560  
Db 181 ACATCCNGTGGAGTGGAGCAGAGCCGACAGGAGGTGGTTCGGTCTGTCTTAGANGCCTC 240  
QY 561 AGGCAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGC-TGGAAGGGATGAAC 619  
Db 241 ANGCAAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGCTTGGAAGGGATGAAC 300  
QY 620 AG-CCAAGTAGTTCAGCTT 639  
Db 301 AGCCCAAGTAGTTCANCTT 321

RESULT 10

US-10-154-884B-3344  
Sequence 3344, Application US/10154884B  
Publication No. US20040005561A1  
GENERAL INFORMATION:  
APPLICANT: Gaiger, Alexander  
APPLICANT: Algarte, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc W.  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
FILE REFERENCE: 014058-013521US  
CURRENT APPLICATION NUMBER: US/10/154,884B  
CURRENT FILING DATE: 2002-05-23  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: US 60/200,303  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,779  
PRIOR FILING DATE: 2000-04-28  
PRIOR APPLICATION NUMBER: US 60/200,999  
PRIOR FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: US 60/202,084  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 60/206,201  
PRIOR FILING DATE: 2000-05-22  
PRIOR APPLICATION NUMBER: US 60/218,950  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: US 60/222,903  
PRIOR FILING DATE: 2000-08-03  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 11290  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3344  
LENGTH: 352  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(352)  
OTHER INFORMATION: n = g, a, c or t  
US-10-154-884B-3344

Query Match 20.5%; Score 256.4; DB 17; Length 352;  
Best Local Similarity 97.2%; Pred. No. 5.1e-67;  
Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

QY 324 CCAGCCGAGCAGCAGCAAGCAACCAAAATTCGTGGGAGAGCAAAAGCTGAGGTCC 383  
Db 1 CCAGCCGAGCAGCAGCAAGCAACCAAAATTCGTGGGAGAGCAAAAGCTGAGGTCC 60  
QY 384 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 443  
Db 61 AGATGAAAGAGAAGTTCTCGTCACTCCCGAGGATTATGCGCGATTCAAGAAATCCATGG 120  
QY 444 TGAAGGTTTGGATTCTTATGAAGAAAAGGAGGATAAAGTGATCAAGGAGATGGCAGCTC 503  
Db 121 TGAAGGTTTGGATTCTTATGAAGAAAAGGAGGATAAAGTGATCAAGGAGATGGCAGCTC 180  
QY 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGGAGGTGGTTCGGTCTGTCTTAGA-GCCTC 560  
Db 181 ACATCCNGTGGAGTGGAGCAGAGCCGACAGGAGGTGGTTCGGTCTGTCTTAGANGCCTC 240  
QY 561 AGGCAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGC-TGGAAGGGATGAAC 619  
Db 241 ANGCAAGTGCAGACCCAGAGAGGGCTCTTCAGCACCTTAGAAGCTTGGAAGGGATGAAC 300  
QY 620 AG-CCAAGTAGTTCAGCTT 639

Db 301 AGCCCAAGTAGTCCANCCTT 321

RESULT 11

US-10-764-324-3344

Sequence 3344, Application US/10764324

Publication No. US20040175739A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

FILE REFERENCE: 014058-013520US

CURRENT APPLICATION NUMBER: US/10/764,324

CURRENT FILING DATE: 2004-01-23

PRIOR APPLICATION NUMBER: US/10/040,862

PRIOR FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: US 60/202,084

PRIOR FILING DATE: 2000-05-04

PRIOR APPLICATION NUMBER: US 60/206,201

PRIOR FILING DATE: 2000-05-22

PRIOR APPLICATION NUMBER: US 60/218,950

PRIOR FILING DATE: 2000-07-14

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 10467

SOFTWARE: FastSEQ for Windows Version 3.0

SEQ ID NO 3344

LENGTH: 352

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: (92)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (187)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (235)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (242)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (318)

OTHER INFORMATION: n=A,T,C or G

FEATURE:

NAME/KEY: unsure

LOCATION: (329)

OTHER INFORMATION: n=A,T,C or G

US-10-764-324-3344

Query Match 20.5%; Score 256.4; DB 19; Length 352;

Best Local Similarity 97.2%; Pred. No. 5.1e-67;

Matches 312; Conservative 0; Mismatches 4; Indels 5; Gaps 5;

Qy 324 CCAGCCAGAGCACAAGAGCAACCAAAATTTCTGGTGGGAGAAACAAAGCTCAGAGTCC 383

Db 1 CCAGCCAGAGCACAAGAGCAACCAAAATTTCTGGTGGGAGAAACAAAGCTCAGAGTCC 60

Qy 384 AGATGAAGAGAGTCTTCTGTCATCTCCCAAGGATTATGCCGATTCAAGAAATCCATGG 443

Db 61 AGATGAAGAGAGTCTTCTGTCATCTCCCAAGGATTATGCCGATTCAAGAAATCCATGG 120

Qy 444 TGAAGGTTTGGATTCTTATGAAGAAAGGAGGATAAGTCAATCAAGGAGATGGCAGCTC 503

Db 121 TGAAGGTTTGGATTCTTATGAAGAAAGGAGGATAAGTCAATCAAGGAGATGGCAGCTC 180

Qy 504 AGATCC-GTGAGGTGGAGCAGAG-CCGACAGAGGTGGTTCGGTCTGTCTTTAGAG-GCCTC 560

Db 181 ACATCCNGTGGAGTGGAGCAGAGCCGACAGGAGTGGTTCGGTCTGTCTTTAGANGCCTC 240

Qy 561 AGCAGTGCACAGCCCAAGAGAGGGCTCTTCAGCACCTAGAGGC-TGAAAAGGGATGAAC 619

Db 241 ANGCAGTGCCAGACCCCAAGAGAGGGCTCTTCAGCACCTAGAGGC-TGAAAAGGGATGAAC 300

Qy 620 AG-CCAAAGTAGTCCAGCTT 639

Db 301 AGCCCAAGTAGTCCANCCTT 321

RESULT 12

US-10-027-632-120356

Sequence 120356, Application US/10027632

Publication No. US20020198371A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 120356

LENGTH: 970

TYPE: DNA

ORGANISM: Human

US-10-027-632-120356

Query Match 17.5%; Score 219.6; DB 13; Length 970;

Best Local Similarity 81.8%; Pred. No. 1.5e-55;

Matches 332; Conservative 1; Mismatches 50; Indels 23; Gaps 6;

Qy 801 CTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAGAAAGGAAAAACAGA 860

Db 3 CTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCTTAAGAAAGGAAAAACAGA 62

Qy 861 AGTTGAAAAAATCTCCCAAGAGCCGAGTCTGGGCCAACTTTTGATCAGAGTCCAGGACCA 920

Db 63 AGTTGAAAGGAACATTTTCCCAAACTTTGAGTGGGGCCAACTTTAATCAGAGTCCAGGACCA 122

Qy 921 GTGCAGGCTGGCTGCCCTCTTTTGGGGCCGGTCTGGAATAATGACGCCGCTGCAGTCC 980



;  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
;  
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES

FILE REFERENCE: 031896-043000 (AM 101081)

; CURRENT APPLICATION NUMBER: US/10/956,157

CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805

; SOFTWARE: Patch

; SOFWAKE: FAW  
; SEO ID NO 902

: LENGTH: 56

```

; LENGTH: 367
; TYPE: DNA

```

TYPE: DNA  
ORGANISM: Homo

Query Match 16.4%; Score 205; DB 21; Length 567;

Best Local Similarity 79.9%; Pred. No. 3e-51;

Accession Number	Similarity	Accession No.	Matches	Conservative	Mismatches	Indels	Gaps
2000	298	0	298	0	55	20	4

QY 411 CCCAGGATTATGCGCGATTCAAGAAATCCATGGTGAAAGGTTGGATTCCCTATGAAGAAA 470

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

QY 471 AGGAGGATAAG--TGATCAAGGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCG 528

Db 256 TAGAGGATTGAGGTGGACCAAGGAGATGGCAGCTCAGATCTGTGAGGTGGTACAGAGCTG 315

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

[illegible]

D<sub>b</sub> 316 GCAGGAGGTGGTTTCAGTCTGTCTCTAGAGGTTGGTTTCCCTCGGAGGATCCAGACCCCTC 375

QY 577 ----AGAAGAGGGCTCTTTCAGCACCTAGAAAGCTGGAAGGATGAACAGCCAAGTAGCT 631

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

QY 632 TCCAGCTTACAGCAGCCCTCAAAATTGGACCTGCCACCAGCTCCAGAGCTTGACTGGATG 691

Db  
436 TCCAGCCCAAGCAGGCCTCCAACTTGGACCTGCCA-CAGCTCCAGAGCTTGACTGGATG 494

Search completed: June 11, 2005, 06:13:08

Job time : 478.179 secs

**THIS PAGE BLANK (USPTO)**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	1252	99.9	1253	6	US-09-155-676B-4	Sequence 4, Appli
2	1252	99.9	1253	13	US-11-060-756-1553	Sequence 1553, A
3	1252	99.9	1253	13	US-11-060-756-1554	Sequence 1554, A
4	1252	99.9	1253	13	US-11-060-756-5825	Sequence 5825, A
5	1252	99.9	1253	13	US-11-060-756-5826	Sequence 5826, A
6	514.6	41.1	681	8	US-10-450-763-19571	Sequence 19571, A
7	479.6	38.3	2557	8	US-10-450-763-23046	Sequence 23046, A
8	370.4	29.6	742	8	US-10-450-763-19570	Sequence 19570, A
9	323.6	25.8	743	15	US-60-680-544-781	Sequence 781, App
C 10	323.6	25.8	743	15	US-60-680-544-7415	Sequence 7415, A
11	323.6	25.8	743	15	US-60-680-473-781	Sequence 781, App
C 12	323.6	25.8	743	15	US-60-680-473-7415	Sequence 7415, A
13	278.4	22.2	816	11	US-10-940-774A-3437	Sequence 3437, A
14	263.4	21.0	3100	8	US-10-450-763-13671	Sequence 13671, A
15	215.2	17.2	4017	8	US-10-450-763-19572	Sequence 19572, A
C 16	175.8	14.0	6623	11	US-10-940-774A-16221	Sequence 16221, A
17	175.8	14.0	7031	11	US-10-940-774A-15179	Sequence 15179, A
18	175.8	14.0	246960	12	US-11-121-086-8	Sequence 8, Appli

Db 61 GATTCGAGCCACGAAGGCCCTTCTCTGTGGTCCGCGACGTTTACAGCCGCAAGCAC 120  
Qy 121 CCAGCGGAGCTGAAGGAGCTTTTCAGAGGCTCCTGCCCCAGAGTGGAGCGGCCGCA 180  
Db 121 CCAGCGGAGCTGAAGGAGCTTTTCAGAGGCTCCTGCCCCAGAGTGGAGCGGCCGCA 180  
Qy 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGCAACACGAGCGATGCTGCTGGTG 240  
Db 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGCAACACGAGCGATGCTGCTGGTG 240  
Qy 241 CTGTGCTGCGGCTGTGAGGTGCGGAAACACTGAGCCATGAAACCTGACGGTCTGTA 300  
Db 241 CTGTGCTGCGGCTGTGAGGTGCGGAAACACTGAGCCATGAAACCTGACGGTCTGTA 300  
Qy 301 CGGGGGCTGCTGGAGCTCTGGCAGCCGACAGCACAGAACCAACCAAAATTCG 360  
Db 301 CGGGGGCTGCTGGAGCTCTGGCAGCCGACAGCACAGAACCAACCAAAATTCG 360  
Qy 361 GTGGAGAAACAAAGCTGAGGTCCAGATGAAAGAGATTTCTGCTCACTCCCGAGATTA 420  
Db 361 GTGGAGAAACAAAGCTGAGGTCCAGATGAAAGAGATTTCTGCTCACTCCCGAGATTA 420  
Qy 421 TGGCGATTCAGAAATCATCGTGGAAGTTTGGATTCCTATGAAGAAAGAGATTA 480  
Db 421 TGGCGATTCAGAAATCATCGTGGAAGTTTGGATTCCTATGAAGAAAGAGATTA 480  
Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGAGTGGAGCAGAGCCGACAGAGGTGGT 540  
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGAGTGGAGCAGAGCCGACAGAGGTGGT 540  
Qy 541 TCGGCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600  
Db 541 TCGGCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600  
Qy 601 AAGCTGGAAGGATGAACAGCAAGTAGTTCAGCTTACAGCAGCCCTCAAAATTTGA 660  
Db 601 AAGCTGGAAGGATGAACAGCAAGTAGTTCAGCTTACAGCAGCCCTCAAAATTTGA 660  
Qy 661 CTGCGCACCAGCTCCAGAGCTTGAAGTGGAGCAGGACCATCTTCGACATTCATGG 720  
Db 661 CTGCGCACCAGCTCCAGAGCTTGAAGTGGAGCAGGACCATCTTCGACATTCATGG 720  
Qy 721 CCATCAGGATATACAGAGTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780  
Db 721 CCATCAGGATATACAGAGTGGTAAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780  
Qy 781 CCAAGATGAAGATACATTTGCTGGAAACCAAGAAATAGGACCATCCTATGAAGAAATTC 840  
Db 781 CCAAGATGAAGATACATTTGCTGGAAACCAAGAAATAGGACCATCCTATGAAGAAATTC 840  
Qy 841 TAAAGAAAGGAAACAGAGCTTGAAGAACTCCCGCAGACGAGTTGGGCGCAACTT 900  
Db 841 TAAAGAAAGGAAACAGAGCTTGAAGAACTCCCGCAGACGAGTTGGGCGCAACTT 900  
Qy 901 TGATCACAGCTCCAGACCAAGTGCAGGCTGGCTGCTCTTTTGGCGCGCTCGGAATA 960  
Db 901 TGATCACAGCTCCAGACCAAGTGCAGGCTGGCTGCTCTTTTGGCGCGCTCGGAATA 960  
Qy 961 ATGAGCGCGCTGGAGTCCAGACATCAACTCCAAAACCTGAAAGCTGCAGCAATGAAGA 1020  
Db 961 ATGAGCGCGCTGGAGTCCAGACATCAACTCCAAAACCTGAAAGCTGCAGCAATGAAGA 1020  
Qy 1021 CAGTCACATACAGAAAGCTTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080  
Db 1021 CAGTCACATACAGAAAGCTTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080  
Qy 1081 AAAGTCAACCAACCCCTATTATACCTTCACCCAAATTCCTTATCATTTCTTTAG 1140  
Db 1081 AAAGTCAACCAACCCCTATTATACCTTCACCCAAATTCCTTATCATTTCTTTAG 1140  
Qy 1141 GAAACAGACATACATTCATTTGATTTAATAAGTTTATTTTTCGGCCTTCGTGGCCT 1200  
Db 1141 GAAACAGACATACATTCATTTGATTTAATAAGTTTATTTTTCGGCCTTCGTGGCCT 1200

Db 1141 GAAACAGACATACATTCATTTGATTTAATAAGTTTATTTTTCGGCCTTCGTGGCCT 1200  
Qy 1201 CGAATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGCGCGTACCCACTTTT 1253  
Db 1201 CGAATCAAGCTTATCGATACCGTGCAGCTCGAGGGGGCGCGTACCCACTTTT 1253  
RESULT 2  
US-11-060-756-1553  
; Sequence 1553, Application US/11060756  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William Martin  
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug  
; TITLE OF INVENTION: Target Genes  
; FILE REFERENCE: AM101083 (031896-042000)  
; CURRENT APPLICATION NUMBER: US/11/060,756  
; CURRENT FILING DATE: 2005-02-18  
; NUMBER OF SEQ ID NOS: 303284  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1553  
; LENGTH: 1253  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (53)..(53)  
; OTHER INFORMATION: n is a, c, g, or t  
US-11-060-756-1553  
Query Match 99.9%; Score 1252; DB 13; Length 1253;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CATTGGAGTACGCGGTGCGCGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60  
Db 1 CATTGGAGTACGCGGTGCGCGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60  
Qy 61 GATTGAGCCACGAAGGCCCTTCTCTGTGGTTCGCGGCACGTTTACAGCCGCAAGCAC 120  
Db 61 GATTGAGCCACGAAGGCCCTTCTCTGTGGTTCGCGGCACGTTTACAGCCGCAAGCAC 120  
Qy 121 CCAGCGCAGCTGAAGGAGGCTTTTGAGAGCTCCTGCCCCAGGTGGAGGCGGCCGCAA 180  
Db 121 CCAGCGCAGCTGAAGGAGGCTTTTGAGAGCTCCTGCCCCAGGTGGAGGCGGCCGCAA 180  
Qy 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGCAACACGAGCGATGCTGCTGGTG 240  
Db 181 GGCATCCGCGCGCTCAGGTGGAGCGCTATGTGCCGCAACACGAGCGATGCTGCTGGTG 240  
Qy 241 CTGTGCTGCGGCTGTGAGGTGCGGAAACACTGAGCCATGAAACCTGACGGTCTGTA 300  
Db 241 CTGTGCTGCGGCTGTGAGGTGCGGAAACACTGAGCCATGAAACCTGACGGTCTGTA 300  
Qy 301 CGGGGGCTGCTGGAGCATCTGGCAGCCGACAGCACAGAACCAACCAAAATTCG 360  
Db 301 CGGGGGCTGCTGGAGCATCTGGCAGCCGACAGCACAGAACCAACCAAAATTCG 360  
Qy 361 GTGGAGAAACAAAGCTGAGGTCCAGATGAAAGAGATTTCTGCTCACTCCCGAGATTA 420  
Db 361 GTGGAGAAACAAAGCTGAGGTCCAGATGAAAGAGATTTCTGCTCACTCCCGAGATTA 420  
Qy 421 TGGCGATTCAGAAATCATCGTGGAAGTTTGGATTCCTATGAAGAAAGAGATTA 480  
Db 421 TGGCGATTCAGAAATCATCGTGGAAGTTTGGATTCCTATGAAGAAAGAGATTA 480  
Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGAGTGGAGCAGAGCCGACAGAGGTGGT 540  
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGAGTGGAGCAGAGCCGACAGAGGTGGT 540  
Qy 541 TCGGCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600  
Db 541 TCGGCTGTCTTAGAGCTCAGGAGTGCAGACCCAGAGAGGCTCTTCAGACCTAG 600



```
QY 601 AAGCTGGAAGGGGATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA 660
Db 601 AAGCTGGAAGGGGATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA 660
QY 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTGACATTCATTGG 720
Db 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTGACATTCATTGG 720
QY 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
QY 781 CCAAGATGAAGATACATCTGCTGGGAACCAAGAAATAGGACATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATACATCTGCTGGGAACCAAGAAATAGGACATCTATGAAGAAATTTCT 840
QY 841 TAAAGAAAGGAAACAGAAAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAACAGAAAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
QY 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAATA 960
Db 901 TGATCACAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAATA 960
QY 961 ATGGACGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGCAGCAATGAAGAAG 1020
Db 961 ATGGACGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGCAGCAATGAAGAAG 1020
QY 1021 CAGTCACATACAGAAAAAGCTATCATGCTCTTACCAACTACCATGAGCTAAAAGCC 1080
Db 1021 CAGTCACATACAGAAAAAGCTATCATGCTCTTACCAACTACCATGAGCTAAAAGCC 1080
QY 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTAG 1140
Db 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTAG 1140
QY 1141 GAAACAGACATACATCATTTGATTTAATAAGTTTTATTTTTCGGCCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATACATCATTTGATTTAATAAGTTTTATTTTTCGGCCCTTCGTGGCCT 1200
QY 1201 CGAATCAGCTTATCGATACCGTCAGCTCGACCTCGAGGGGGCGGTACCCACTTTT 1253
Db 1201 CGAATCAGCTTATCGATACCGTCAGCTCGACCTCGAGGGGGCGGTACCCACTTTT 1253
```

## RESULT 3

```
US-11-060-756-1554
; Sequence 1554, Application US/11060756
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TITLE OF INVENTION: Target Genes
; FILE REFERENCE: AM101083 (031896-042000)
; CURRENT APPLICATION NUMBER: US/11/060,756
; CURRENT FILING DATE: 2005-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1554
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-11-060-756-1554
```

```
Query Match 99.9%; Score 1252; DB 13; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CATTGGAGTCACGCGGTGGCGCGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
Db 1 CATTGGAGTCACGCGGTGGCGCGCTCTAGAAATAGTGGATCCCGGGCTGCANGGAATTC 60
QY 61 GATTTCGAGCCACCAAGGCCCTTTCTTCTGTGGTTCGGGACAGTTTACAGCCGCAAGCAC 120
Db 61 GATTTCGAGCCACCAAGGCCCTTTCTTCTGTGGTTCGGGACAGTTTACAGCCGCAAGCAC 120
QY 121 CCAGCGCAGCTGAAGAGGCGCTTTTGAGAGGCTCTCTGCCCCAGGTGGAGGGCGGCCGCA 180
Db 121 CCAGCGCAGCTGAAGAGGCGCTTTTGAGAGGCTCTCTGCCCCAGGTGGAGGGCGGCCGCA 180
QY 181 GGCCATCGCGCCCTCAGGTGGAGCGCTATGTGCCCGAAACAGGAGCATCTCTGTGTG 240
Db 181 GGCCATCGCGCCCTCAGGTGGAGCGCTATGTGCCCGAAACAGGAGCATCTCTGTGTG 240
QY 241 CTTGTGCTGCGGCTGTGAGGTCGCGGAAACACTCAGGCCATCGAAACCTTGACGGTCTGTA 300
Db 241 CTTGTGCTGCGGCTGTGAGGTCGCGGAAACACTCAGGCCATCGAAACCTTGACGGTCTGTA 300
QY 301 CGGGGGCTGTCTGGAGCATCTGGCCAGCCAGAGCAACAAGAAAGCAACAATTTCTG 360
Db 301 CGGGGGCTGTCTGGAGCATCTGGCCAGCCAGAGCAACAAGAAAGCAACAATTTCTG 360
QY 361 GTGGGAGAACAAAGCTGAGGTCCAGATGAAAGAGAGTTTCTGTCTCACTCCCAAGATTA 420
Db 361 GTGGGAGAACAAAGCTGAGGTCCAGATGAAAGAGAGTTTCTGTCTCACTCCCAAGATTA 420
QY 421 TGCCCGATTCAAGAAATCCATGTGTGAAGTTTCGATTTCCTATGAAGAAAGGAGGATAA 480
Db 421 TGCCCGATTCAAGAAATCCATGTGTGAAGTTTCGATTTCCTATGAAGAAAGGAGGATAA 480
QY 481 AGTGATCAAGAGATGCGAGCTCAGATCCGTGTGGTGGAGCAGAGCCGACAGGAGTGT 540
Db 481 AGTGATCAAGAGATGCGAGCTCAGATCCGTGTGGTGGAGCAGAGCCGACAGGAGTGT 540
QY 541 TCGTCTGTCTTAGAGCTCAGGAGTCCAGACCCAGAGAGGGCTCTTTCAGCACTTAG 600
Db 541 TCGTCTGTCTTAGAGCTCAGGAGTCCAGACCCAGAGAGGGCTCTTTCAGCACTTAG 600
QY 601 AAGCTGGAAGGGGATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA 660
Db 601 AAGCTGGAAGGGGATGAACAGCCAAAGTAGCTTCCAGCTTACAGCAGCCCTCAAATTTGGA 660
QY 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720
Db 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720
QY 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACCAAGGATTTGTAACATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
QY 781 CCAAGATGAAGATACATCTGCTGGGAACCAAGAAATAGGACATCTTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATACATCTGCTGGGAACCAAGAAATAGGACATCTTATGAAGAAATTTCT 840
QY 841 TAAAGAAAGGAAACAGAAAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAACAGAAAGTTGAAAGAACTCCCCCCAGACCGAGTTGGGGCCAACTT 900
QY 901 TGATCAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAATA 960
Db 901 TGATCAGCTCCAGGACCAAGTGCAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTGGAATA 960
QY 961 ATGGACGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGCAGCAATGAAGAAG 1020
Db 961 ATGGACGCGCTGGCAGTCCAGACATCAATCCAAACTGAAAGCTGCAGCAATGAAGAAG 1020
QY 1021 CAGTCACATACAGAAAAAGCTATCATGCTCTTACCAACTACCATGAGCTAAAAGCC 1080
Db 1021 CAGTCACATACAGAAAAAGCTATCATGCTCTTACCAACTACCATGAGCTAAAAGCC 1080
QY 1081 AAAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTTCTAG 1140
```



Query Match 99.9%; Score 1252; DB 13; Length 1253;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGANGGAATTC 60  
Db 1 CATTGGAGTCACGGGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGANGGAATTC 60

Qy 61 GATTTCAGCCCAAGAGGCTCTCTCTGGTTCGGGCACTGTTACAGCCGCAAGCAC 120  
Db 61 GATTTCAGCCCAAGAGGCTCTCTCTGGTTCGGGCACTGTTACAGCCGCAAGCAC 120

Qy 121 CCAGCGCAGCTGAAGAGGCTTTTGAAGAGCTCTCTGCCCCAGGTGGAGGCGGCCGCAA 180  
Db 121 CCAGCGCAGCTGAAGAGGCTTTTGAAGAGCTCTCTGCCCCAGGTGGAGGCGGCCGCAA 180

Qy 181 GGCCATCCGCGCTCAGGTGGAGGCTATGTGCCGAACACGAGGATCTCTGGTG 240  
Db 181 GGCCATCCGCGCTCAGGTGGAGGCTATGTGCCGAACACGAGGATCTCTGGTG 240

Qy 241 CCTGTGCTGCGGCTGTAGGTGGCGGAACACTGAGCCATCGAAACCTGACGGTGTGTA 300  
Db 241 CCTGTGCTGCGGCTGTAGGTGGCGGAACACTGAGCCATCGAAACCTGACGGTGTGTA 300

Qy 301 CGGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTG 360  
Db 301 CGGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTG 360

Qy 361 GTGGAGNACAAAGCTGAGGTCAGATCAAGAGAGTTTCTGTCACTCCCAAGGATTA 420  
Db 361 GTGGAGNACAAAGCTGAGGTCAGATCAAGAGAGTTTCTGTCACTCCCAAGGATTA 420

Qy 421 TGCCGATTCAAGAAATCCATGGTCAAGGTTTGGATTCTTATCAAGAAAGGAGGATAA 480  
Db 421 TGCCGATTCAAGAAATCCATGGTCAAGGTTTGGATTCTTATCAAGAAAGGAGGATAA 480

Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGTGT 540  
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGGAGTGT 540

Qy 541 TCAGTCTGTCTTAGAGCTCAGGAGTGCAGAGCCAGAGAGGAGGCTTTCAGACCTAG 600  
Db 541 TCAGTCTGTCTTAGAGCTCAGGAGTGCAGAGCCAGAGAGGAGGCTTTCAGACCTAG 600

Qy 601 AAGCTGGAAGAGGATGAACAGCCAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGA 660  
Db 601 AAGCTGGAAGAGGATGAACAGCCAGTAGCTTCCAGCTTACAGAGCCCTCAAAATTTGA 660

Qy 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720  
Db 661 CCTGCCACAGCTCCAGAGCTTGACTGGATGGAGACAGGACCATCTCTGACATTCATTGG 720

Qy 721 CCATCAGATATACAGAGTTGGTAAATCACTCAGGTGGCCACACTCCTCCTGGATGAT 780  
Db 721 CCATCAGATATACAGAGTTGGTAAATCACTCAGGTGGCCACACTCCTCCTGGATGAT 780

Qy 781 CCAAGATGAAGATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840  
Db 781 CCAAGATGAAGATACATTTGCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840

Qy 841 TAAAGAAAGGAAAAACAGAAATTTGAAAAAACTCCCGCCAGAGCTGGGGCCAACTT 900  
Db 841 TAAAGAAAGGAAAAACAGAAATTTGAAAAAACTCCCGCCAGAGCTGGGGCCAACTT 900

Qy 901 TGATCAGAGTCCAGGACCAAGTGCAGGTGGCTGCCCTCTTTTGGGCGCGTCTGGAATA 960  
Db 901 TGATCAGAGTCCAGGACCAAGTGCAGGTGGCTGCCCTCTTTTGGGCGCGTCTGGAATA 960

Qy 961 ATGAGCCGCTGCGAGTCCAGACATCACTCCAAAATGAGCTGCAGCATGAAGAAG 1020  
Db 961 ATGAGCCGCTGCGAGTCCAGACATCACTCCAAAATGAGCTGCAGCATGAAGAAG 1020

RESULT 6

US-10-450-763-19571  
; Sequence 19571, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 19571  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (187)..(675)  
; OTHER INFORMATION: 78% homologous to Homo sapiens TRAF2 binding protein encoded  
; OTHER INFORMATION: by clone 15, accession number W42401, Smith-Waterman Score=665.  
US-10-450-763-19571

Query Match 41.1%; Score 514.6; DB 8; Length 681;  
Best Local Similarity 97.4%; Pred. No. 8.6e-124;  
Matches 568; Conservative 0; Mismatches 4; Indels 11; Gaps 4;

Qy 81 CCTTCTTCTGTGTGGTGGC-GGCACGTTTACAGCCGCAAGCACCAGCGGACGCTGAAGGAG 139  
Db 99 CCTTCTTCTGTGTGGTGGCAGCGTTCAGCGCCGCAAGCA-CCAGCGGACGCTGAAGGAG 157

Qy 140 GCTTTTGAAGAGGCTCTCTGCCCC-----AGTGGAGGCGGCGCCGCAAGCCATCCGCG 191  
Db 158 GCTTTTGAAGAGGCTCTCTGCCCCAGGTGGAGGCGGCGCCGCAAGCCATCCGCG 217

Qy 192 CCCTCAGGTGGAGCGCTATGTGCCGAAACAGCGCGATGCTGCTGCTGCTGCTGCTGCTGCG 251  
Db 218 CCCTCAGGTGGAGCGCTATGTGCCGAAACAGCGCGATGCTGCTGCTGCTGCTGCTGCTGCG 277

Qy 252 GCTGTGAGGTGGCGGAACACCTGAGCCATGGAACCTGACGCTGCTGTACGGGGGGTGC 311  
Db 278 GCTGTGAGGTGGCGGAACACCTGAGCCATGGAACCTGACGCTGCTGTACGGGGGGTGC 337

Qy 312 TGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTTGTGGGAGAAC 371  
Db 338 TGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATCTTGTGGGAGAAC 397

Qy 372 AAGCTGAGGTCCAGATGAAGAGAGTTTCTGTGCTCACTCCCAAGGATTTGCGCGATTCA 431  
Db 398 AAGCTGAGGTCCAGATGAAGAGAGTTTCTGTGCTCACTCCCAAGGATTTGCGCGATTCA 457

QY 432 AGAATCCATGTTGAAAGGTTGGATTCTCTATGAAGAAAGAGAGATAAAGTGATCAAGG 491  
Db 458 AGAAATCCATGTTGAAAGGTTGGATTCTCTATGAAGAAAGAGAGATAAAGTGATCAAGG 517  
QY 492 AGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACA-GGAGGTGGTTTGGTCTGTC 550  
Db 518 AGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACA-GGAGGTGGTTTGGTCTGTC 577  
QY 551 TTAGAGCCTCAGGAGTGCAGACCCAGAGAGGGCTCTTACAGCACCTAGAAAGCTGGAAA 610  
Db 578 TTAGAGCCTCAGGAGTGCAGACCCAGAGAGGGCTCTTACAGCACCTATTAAGCTGGAAA 637  
QY 611 GGGATGAACAGCAAGTACGTTCCAGCTTACAGCAGCCCTCAA 653  
Db 638 GGGATGAACAGCAAGTACGTTCCAGCTTACAGCAGCCCTTAA 680

## RESULT 7

US-10-450-763-23046  
; Sequence 23046, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 23046  
; LENGTH: 2557  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1885)..(2196)  
; OTHER INFORMATION: 40% homologous to Mus musculus synbindin, accession number AF233340, Smith-Waterman Score=116.  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2557)  
; OTHER INFORMATION: n = a,t,c or g  
US-10-450-763-23046

Query Match 38.3%; Score 479.6; DB 8; Length 2557;  
Best Local Similarity 84.9%; Pred. No. 1.8e-114;  
Matches 702; Conservative 0; Mismatches 55; Indels 70; Gaps 12;

QY 426 GATTCAGAAATCCATGTTGAAAGGTTGGATTCTCTATGAAGAAAGAGAGATAAAGTGA 485  
Db 36 GATTCAGAAATCCATGTTGAAAGGTTGGATTCTCTATGAAGAAAGAGAGATAAAGTGA 95  
QY 486 TCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGGTTGTCGGT 545  
Db 96 TCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGAGCCGACAGAGGTTGTCGGT 155  
QY 546 CTGTCTTTAGAG-----CCTCAGGCGAGTGCAGACCCA 577  
Db 156 CTGTCTTTAGAGTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGCGAGTGCAGACCCA 215  
QY 578 GAAGAGGGCTCTTACAGCAGCTA-GAAGCTGGAAGAGGATGAACAGCCAGTAGCTTT-CCA 635  
Db 216 GAAGAGGGCTCTTACAGCAGCTAAGAGCTGGAAGGATGAACAGCCAGTAGCTTTCCA 275  
QY 636 GCTTACAGCAGCCCTCAAATTTGGACCTGCCACACAGCTCCAGAGCTTCACTGGATGGAGA 695  
Db 276 GCTTACAGCAGCCCTCAAATTTGGACCTGCCACACAGCTCCAGAGCTTCACTGGATGGAGA 335

QY 696 CAGGACCATCTCTGATTCATTTGTCCTATGTCCTATGATATATA-CCAGGAGTTGGTAAACATCCAC 754  
Db 336 CAGGACCATCTCTGATTCATTTGTCCTATGTCCTATGATATATA-CCAGGAGTTGGTAAACATCCAC 395  
QY 755 TCAGGTGCCACACCTCTCCCTGGATGATCCAAAGATGAAGATACAT-----TGCTGGGAACCA 810  
Db 396 TCAGGTGCCACACCTCTCCCTGGATGATCCAAAGATGAAGATACATTTGGCTGGGGGACCCCA 455  
QY 811 AGAAATAGGACC--ATCCTATGAAGAATTTCTTTAAAGAAA----- 848  
Db 456 AGAAATNGGGCCCATCTCTATGAAGAATTTCTTTAAAGAAATTTCCCTCTATCATTTGTAC 515  
QY 849 --AGGAAAAACAGAGTTGAAAAAACTCCCCC-----AGACCGAGTTGGGG--CCAACCTT 900  
Db 516 AGAGGAAAAACAGAGTTGAAAAAACTCCCCC-----AGACCGAGTTGGGG--CCAACCTT 575  
QY 901 TGATCAGAGTCC--AGGACGAGTGCAGGCTGGCTGGCTCTCTTTTGGGCCCGGCTCTGGAA 958  
Db 576 GGATCAGAGTCCCGAGGACCCAGTGCAGGCTGGCTGGCTCTCTTTTGGGCCCGGCTCTGGAA 635  
QY 959 TAATGGACGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGA 1018  
Db 636 AATCGAGCGCGCTGGCAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGA 695  
QY 1019 AGCAGTCAATACAGAAAAAGAGTAAATCATGCTCTCTACCAACTACCATGAGGCTAAAG 1078  
Db 696 AGCAGTCAATACAGAAAAAGAGTAAATCATGCTCTCTACCAACTACCATGAGGCTAAAG 755  
QY 1079 CAAAAGTCAACAAACCCCTATTATACCTTCACCCAAATTTCTTTATCATTTCTTTCTT 1138  
Db 756 C--AAAGTCAACAAACCCCTATTATACCTTCCA-CCAAATTTCTTTATCATTTCTTTCTT 812  
QY 1139 AGGAAACAGACATCTCTCATTTGATTTAAATAAGTTTATTTT 1185  
Db 813 AGGAAACAGACATCTCTCATTTGATTTAAATAAGTTTAAATTT 859

## RESULT 8

US-10-450-763-19570  
; Sequence 19570, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 19570  
; LENGTH: 742  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (307)..(642)  
; OTHER INFORMATION: 67% homologous to Homo sapiens TRAF2 binding protein encoded by clone 15, accession number W42401, Smith-Waterman Score=349.  
US-10-450-763-19570

Query Match 29.6%; Score 370.4; DB 8; Length 742;  
Best Local Similarity 97.4%; Pred. No. 3.5e-86;  
Matches 419; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

QY 757 AGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTCTGGTGGGACCAAGAAAT 816  
Db 306 AGGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTTCTGGTGGGACCAAGAAAT 365

Qy	817	AGGACCACTCTATGACGAATTTCTTAAAGAAAAGGAAAAACAGAGTTGTAAGAAAACCTCCC	876
Db	366	AGGACCACTCTATGACGAATTTCTTAAAGAAAAGGAAAAACAGAGTTGTAAGAAAACCTCCC	425
Qy	877	CCAGACCGAGTTGGGGCCAACTTTTGATCACAGTCCAGGACACAGTCGAGCTGGCTGCC	936
Db	426	CCAGACCGAGTTGGGGCCAACTTTTGATCACAGTCCAGGACACAGTCGAGCTGGCTGCC	485
Qy	937	CTCTTTTGGGCGCGTCTGGAATAATGGACGCGCGTGGCAGTCCAGACATCAACTCCAAA	996
Db	486	CTCTTTT-GGCGCGTCTGGAATAATGGACGCGCGTGGCAGTCCAGACATCA-TTCAA	543
Qy	997	ACTGAAGCTCGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTTAATCATGCTCTCTA	1056
Db	544	ACTGAAGCTCGAGCAATGAAGAGCAGTCAATACAGAAAAAGCTTAATCATGCTCTCTA	603
Qy	1057	CCAACTACCATGAGGCTTAAAGCCAAAGTCAACCAACCCCTATTACCTTCCACCCCA	1116
Db	604	CCAACTACCATGAGGCTTAAAGC--AAAGTCAACAAACCCCTATTATACCTTCCA-CCAA	660
Qy	1117	ATTCTTTATCATTTGTTCTTTTAGGAAACAGACATCTCATTTCAATTTGATTTTAATAAAGT	1176
Db	661	ATTCTTTATCATTTGTTCTTTTAGGAAACAGACATCTCATTTCAATTTGATTTTAATAAAGT	720
Qy	1177	TTTTATTTTC	1186
Db	721	TTTTATTTTC	730

```

RESULT 9
US-60-680-544-781
; Sequence 781, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary
; TITLE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 781
; LENGTH: 743
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(743)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-544-781

```

Qy	905	CACAGCTCCAGGACCAAGTCAGGCTGGCTGCCCTCTTTGGCGCGGCTCTGGAATAATGG	964
Db	500	CACAGCTCCAGGACCAAGTCAGGCTGGCTGCCCTCTTTT-GGCCGAGTCTGGAATAATGG	558
Qy	965	ACGCGCTGCGAGTCGAGCATCAATCCAAAACCTGAAGCTGCAGCAATGAAGAAGCAGT	1024
Db	559	ACGACGCTGCGAGTCGAGCATCAA-TTCAAACTGAAGCTGCAGCAATGAAGAAGCAGT	617
Qy	1025	CACATACAGAAAAAGCTTAATCATGCTCTCTACCAACTACCATGAGCTAAAAGCCAAAG	1084
Db	618	CACATACAGAAAAAGCTTAATCATG--CTTTACCAACTACCATGAGCTAAAAG-CANAG	674
Qy	1085	TCAACCAAAACCCCTATTATACCTTTCCACCACAAATCTTTTATCATCTCTTTCTTAGAAA	1144
Db	675	TCAACAAACCCC-----TACTTCCACGGAATCTTTATCATGTCTTTCTTAGAAA	727
Qy	1145	CAGACATACTCATTTCA 1160	
Db	728	CAGACATACTCATTTCA 743	

RESULT 10  
 US-60-680-544-7415/c  
 ; Sequence 7415, Application US/60680544  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cooper, Matthew  
 ; APPLICANT: Kinch, Deborah  
 ; APPLICANT: Rosenberg, Michael  
 ; APPLICANT: Subramaniam, S. Sai  
 ; APPLICANT: Szak, Suzanne  
 ; APPLICANT: Li, Huo  
 ; APPLICANT: Bandaru, Raj  
 ; APPLICANT: Derbel, Maher  
 ; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Compen  
 ; FILE OF INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use There  
 ; FILE REFERENCES: 21590290000  
 ; CURRENT APPLICATION NUMBER: US/60/680,544  
 ; CURRENT FILING DATE: 2005-05-13  
 ; NUMBER OF SEQ ID NOS: 48714  
 ; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
 ; SEQ ID NO 7415  
 ; LENGTH: 743  
 ; TYPE: DNA  
 ; ORGANISM: Macaca Fascicularis  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)...(743)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-60-680-544-7415

Query Match	25.8%;	Score 323.6;	DB 15;	Length 743;
Best Local Similarity	93.3%;	Pred. No. 5.5e-74;		
Matches 407;	Conservative 0;	Mismatches 14;	Indels 15;	Gaps 6;
Qy	725	CAGGATATACCGAGGTTGGTAACTC	CACTCAGGTGGCACACCTCCCTGGATGATCCAA	784
Db	421	CAGGATATACCGAGGTTGGTAACTC	CACTCAGGTGGCACACCTCCCTGGATGATCCAA	352
Qy	785	GATGAAGAATACATTGCTGGGAA	CCAGAATAAGGACCATCTCTATGAAGAATTTCTTAAA	844
Db	361	GAT---GAATACATCGTGGGAAC	CAAGAAATAGGACCATCTATGAAGAATTTCTTAAA	305
Qy	845	GAAGAAGAAAACAGAAAGTTG	AAAAAACTCCCCCAGACCGAGTTGGGGCCCAACTTTGAT	904
Db	304	GAAGAAGAAAACAGAAAGTTG	AAAAAACTCCCCCAGACCGAGTTGGGGCCCAACTTTGAT	245
Qy	905	CACAGCTCCAGGACAGTGCAGCT	GGCTGGCTCTTTTGGCGCGGCTCTCGGAATAATCG	964
Db	244	CACAGCTCCAGGACAGTGCAGCT	GGCTGGCTCTTTTGGCGCGGCTCTCGGAATAATCG	186
Qy	965	ACGCCGCTGGCAGTCCAGACAT	CAACTCCAAAACTGAAAGCTGCAGCAATGAAGAAGCAGT	1024

Db 185 ACAGCCTGGAGTCCAGACATCAA-TTCAAACTGAAGCTGCAGCAATGAAGAAGCAGT 127  
QY 1025 CACATACAGAAAAAGCTAATCATGCTCTTACCACTACCATGAGGCTAAAAGCCAAAG 1084  
Db 126 CACATACAGAAAAAGCTAATCATG--CTTTACCACTACCATGAGGCTAAAAG-CAAAG 70  
QY 1085 TCAACAAAACCCCTATTATACCTTCCACCAAATCTTTATCATTTCTTTCTTAGGAAA 1144  
Db 69 TCAACAAAACCC-----TACTTTCCACCGAATCTTTATCATTTCTTTCTTAGGAAA 17  
QY 1145 CAGACATACCTATTC 1160  
Db 16 CAGACATACCTATTC 1

## RESULT 11

US-60-680-473-781  
; Sequence 781, Application US/60680473  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 781  
; LENGTH: 743  
; TYPE: DNA  
; ORGANISM: Macaca Fascicularis  
; NAME/KEY: misc feature  
; LOCATION: (1)...(743)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-680-473-781

Query Match 25.8%; Score 323.6; DB 15; Length 743;  
Best Local Similarity 93.3%; Pred. No. 5.5e-74;  
Matches 407; Conservative 0; Mismatches 14; Indels 15; Gaps 6;  
QY 725 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 784  
Db 323 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 382  
QY 785 GATGAAGATACATTTGCTGGGAACCAAGAAATAGGACCATCTTATGAAGAATTTCTTAAA 844  
Db 383 GAT---GAATACATCGTGGGAACCAAGAAATAGGACCATCTTATGAAGAATTTCTTAAA 439  
QY 845 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGAT 904  
Db 440 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGAT 499  
QY 905 CACAGCTCCAGGACCAAGTTCGAGTGGCTGGCTCCCTCTTTTGGGCCGCGTCTGGAATAATGG 964  
Db 500 CACAGCTCCAGGACCAAGTTCGAGTGGCTGGCTCCCTCTTTT-GGCCGAGTCTGGAATAATGG 558  
QY 965 ACGCCGCTGGAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCATGAAGAAGCAGT 1024  
Db 559 ACGCCGCTGGAGTCCAGACATCAA-TTCAAACTGAAGCTGCAGCATGAAGAAGCAGT 617  
QY 1025 CACATACAGAAAAAGCTAATCATGCTCTTACCAACTACCATGAGGCTAAAAGCCAAAG 1084  
Db 618 CACATACAGAAAAAGCTAATCATG--CTTTACCACTACCATGAGGCTAAAAG-CAAAG 674  
QY 1085 TCAACAAAACCCCTATTATACCTTCCACCAAATCTTTATCATTTCTTTCTTAGGAAA 1144

Db 675 TCAACAAAACCC-----TACTTTCCACCGAATCTTTATCATTTCTTTCTTAGGAAA 727  
QY 1145 CAGACATACCTATTC 1160  
Db 728 CAGACATACCTATTC 743

## RESULT 12

US-60-680-473-7415/c  
; Sequence 7415, Application US/60680473  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,473  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 7415  
; LENGTH: 743  
; TYPE: DNA  
; ORGANISM: Macaca Fascicularis  
; NAME/KEY: misc feature  
; LOCATION: (1)...(743)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-680-473-7415

Query Match 25.8%; Score 323.6; DB 15; Length 743;  
Best Local Similarity 93.3%; Pred. No. 5.5e-74;  
Matches 407; Conservative 0; Mismatches 14; Indels 15; Gaps 6;  
QY 725 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 784  
Db 421 CAGGATATACAGGAGTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAA 362  
QY 785 GATGAAGATACATTTGCTGGGAACCAAGAAATAGGACCATCTTATGAAGAATTTCTTAAA 844  
Db 361 GAT---GAATACATCGTGGGAACCAAGAAATAGGACCATCTTATGAAGAATTTCTTAAA 305  
QY 845 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGAT 904  
Db 304 GAAAGGAAAAACAGAAGTTGAAAAAACTCCCCCAGACCGAGTTGGGGCCAACTTTGAT 245  
QY 905 CACAGCTCCAGGACCAAGTTCGAGTGGCTGGCTCCCTCTTTTGGGCCGCGTCTGGAATAATGG 964  
Db 244 CACAGCTCCAGGACCAAGTTCGAGTGGCTGGCTCCCTCTTTT-GGCCGAGTCTGGAATAATGG 186  
QY 965 ACGCCGCTGGAGTCCAGACATCACTCCAAAACCTGAAGCTGCAGCATGAAGAAGCAGT 1024  
Db 185 ACAGCCTGGAGTCCAGACATCAA-TTCAAAACCTGAAGCTGCAGCATGAAGAAGCAGT 127  
QY 1025 CACATACAGAAAAAGCTAATCATGCTCTTACCAACTACCATGAGGCTAAAAGCCAAAG 1084  
Db 126 CACATACAGAAAAAGCTAATCATG--CTTTACCACTACCATGAGGCTAAAAG-CAAAG 70  
QY 1085 TCAACAAAACCCCTATTATACCTTCCACCAAATCTTTATCATTTCTTTCTTAGGAAA 1144  
Db 69 TCAACAAAACCC-----TACTTTCCACCGAATCTTTATCATTTCTTTCTTAGGAAA 17  
QY 1145 CAGACATACCTATTC 1160  
Db 16 CAGACATACCTATTC 1

```

RESULT 13
US-10-940-774A-3437
; Sequence 3437, Application US/10940774A
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774A
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3437
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Human
US-10-940-774A-3437

```

Query Match	22.2%;	Score 278.4;	DB 11;	Length 816;
Best Local Similarity	96.7%;	Pred. No. 3.5e-62;		
Matches 327;	Conservative 0;	Mismatches 6;	Indels 5;	Gaps 4;
Qy	849	AGGAAAAACAGAAGTGTGAAAAAATCTCCCCACAGACCGAGTTGGGGCCCAACTTTGATCACA	908	
Db	1	AGGAAAAACAGAAGTGTGAAAAAATCTCCCCACAGACCGAGTTGGGGCCCAACTTTGATCACA	60	
Qy	909	GCTCCAGGACCAAGTCGAGGCTGGCTGCCCTCTTTTGGGCGCGCTCTCGAATAATGAGACGC	968	
Db	61	GCTCCAGGACCAAGTCGAGGCTGGCTGCCCTCTTTT-GGCCGCGCTCGAATAATGAGACGC	119	
Qy	969	CGCTGGCGATCCAGACATCAACTCCAAAACTGAAGCTGCGACGAATGAAGAAGCAGTCA	1028	
Db	120	CGCTGGCGATCCAGACATCAA--TTCAAAACTGAAGCTGCGACGAATGAAGAAGCAGTCA	178	
Qy	1029	TACAGAAAAAGCTAAATCATGCTCTCTACCAACTACCATGAGGCTAAAGCCCAAGTCAA	1088	
Db	179	TACAGAAAAAGCTAAATCATGCTCTCTACCAACTACCATGAGGCTAAAGC--AAAGTCA	236	
Qy	1089	CCAAACCCCTATTATACCTTCCACCCCAAAATCTTTTATCATTTGCTTTCTTAGGAAAACAGA	1148	
Db	237	ACAAACCCCTATTATACCTTCCA--CCAAATCTTTTATCATTTGCTTTCTTAGGAAAACAGA	295	
Qy	1149	CATACTCATTCATTGATTTTAATAAAGTTTTTATTTTC	1186	
Db	296	CATACTCATTCATTGATTTTAATAAAGTTTTTATTTTC	333	

```

RESULT 14
US-10-450-763-13671
; Sequence 13671, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hvaeq. Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 13671
; LENGTH: 3100

```

[illegible]

Mon Jun 13 13:27:14 2005

Db 808 GGTGCCACACCTCCCTGGATGATCCAGATGAAGATACATTGCTGGAAACCAAGAATA 867  
Qy 818 GGACCATCCTATGAAGAATTTCCTTAAGAAAGGAAAAACAGAAAGTTGAAAAAACTCCCC 877  
Db 868 GGACCATCCTATGAAGAATTTCCTTAAGAAAGGAAAAACAGAAAGTTGAAAAAACTCCCC 927  
Qy 878 CCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTCAGGCTGGTGCCC 937  
Db 928 CCAGACCGAGTTGGGGCCAACTTTGATCACAGCTCCAGGACCAAGTCAGGCTGGTGCCC 987  
Qy 938 TCTTTTGGGCGCGTCTGGAATAATGGACGCGCTGGCAGTCCAGACATCAA 989  
Db 988 TCTTTT-GGCCGCGTCTGGAATAATGGACGCGCTGGCAGTCCAGAGCTCCA 1038

Search completed: June 11, 2005, 05:07:14  
Job time : 482.315 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 3133.95 Seconds  
(without alignments)  
16321.674 Million cell updates/sec

Title: US-09-155-676B-4

Perfect score: 1253

Sequence: 1 cattggaggtacgcggtggc.....ggggggcgctaccacatttt 1253

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main.\*  
1: /cgn2\_6/ptodata/1/pna/PTUS1\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/PTUS2\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/PTUS3\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US08\_COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq.\*  
10: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq.\*  
12: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq.\*  
13: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq.\*  
14: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq.\*  
15: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq.\*  
16: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq.\*  
17: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq.\*  
18: /cgn2\_6/ptodata/1/pna/US092A\_COMB.seq.\*  
19: /cgn2\_6/ptodata/1/pna/US092B\_COMB.seq.\*  
20: /cgn2\_6/ptodata/1/pna/US093A\_COMB.seq.\*  
21: /cgn2\_6/ptodata/1/pna/US093B\_COMB.seq.\*  
22: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq.\*  
23: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq.\*  
24: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq.\*  
25: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq.\*  
26: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq.\*  
27: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq.\*  
28: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq.\*  
29: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq.\*  
30: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq.\*  
31: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq.\*  
32: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq.\*  
33: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq.\*  
34: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq.\*  
35: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq.\*  
36: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq.\*  
37: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq.\*  
38: /cgn2\_6/ptodata/1/pna/US098D\_COMB.seq.\*  
39: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq.\*  
40: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq.\*  
41: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq.\*  
42: /cgn2\_6/ptodata/1/pna/US099D\_COMB.seq.\*  
43: /cgn2\_6/ptodata/1/pna/US099E\_COMB.seq.\*

44: /cgn2\_6/ptodata/1/pna/US099F\_COMB.seq.\*  
45: /cgn2\_6/ptodata/1/pna/US099G\_COMB.seq.\*  
46: /cgn2\_6/ptodata/1/pna/US100A\_COMB.seq.\*  
47: /cgn2\_6/ptodata/1/pna/US100B\_COMB.seq.\*  
48: /cgn2\_6/ptodata/1/pna/US101A\_COMB.seq.\*  
49: /cgn2\_6/ptodata/1/pna/US101B\_COMB.seq.\*  
50: /cgn2\_6/ptodata/1/pna/US102A\_COMB.seq.\*  
51: /cgn2\_6/ptodata/1/pna/US102B\_COMB.seq.\*  
52: /cgn2\_6/ptodata/1/pna/US103A\_COMB.seq.\*  
53: /cgn2\_6/ptodata/1/pna/US103B\_COMB.seq.\*  
54: /cgn2\_6/ptodata/1/pna/US104A\_COMB.seq.\*  
55: /cgn2\_6/ptodata/1/pna/US104B\_COMB.seq.\*  
56: /cgn2\_6/ptodata/1/pna/US105A\_COMB.seq.\*  
57: /cgn2\_6/ptodata/1/pna/US105B\_COMB.seq.\*  
58: /cgn2\_6/ptodata/1/pna/US106A\_COMB.seq.\*  
59: /cgn2\_6/ptodata/1/pna/US107A\_COMB.seq.\*  
60: /cgn2\_6/ptodata/1/pna/US107B\_COMB.seq.\*  
61: /cgn2\_6/ptodata/1/pna/US107C\_COMB.seq.\*  
62: /cgn2\_6/ptodata/1/pna/US107D\_COMB.seq.\*  
63: /cgn2\_6/ptodata/1/pna/US108A\_COMB.seq.\*  
64: /cgn2\_6/ptodata/1/pna/US108B\_COMB.seq.\*  
65: /cgn2\_6/ptodata/1/pna/US109A\_COMB.seq.\*  
66: /cgn2\_6/ptodata/1/pna/US109B\_COMB.seq.\*  
67: /cgn2\_6/ptodata/1/pna/US109C\_COMB.seq.\*  
68: /cgn2\_6/ptodata/1/pna/US110\_COMB.seq.\*  
69: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq.\*  
70: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq.\*  
71: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq.\*  
72: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq.\*  
73: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq.\*  
74: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq.\*  
75: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq.\*  
76: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq.\*  
77: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq.\*  
78: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq.\*  
79: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq.\*  
80: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq.\*  
81: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq.\*  
82: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq.\*  
83: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq.\*  
84: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq.\*  
85: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq.\*  
86: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq.\*  
87: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq.\*  
88: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq.\*  
89: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq.\*  
90: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq.\*  
91: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq.\*  
92: /cgn2\_6/ptodata/1/pna/US6023A\_COMB.seq.\*  
93: /cgn2\_6/ptodata/1/pna/US6023B\_COMB.seq.\*  
94: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq.\*  
95: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq.\*  
96: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq.\*  
97: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq.\*  
98: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq.\*  
99: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq.\*  
100: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq.\*  
101: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq.\*  
102: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq.\*  
103: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq.\*  
104: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq.\*  
105: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq.\*  
106: /cgn2\_6/ptodata/1/pna/US6037\_COMB.seq.\*  
107: /cgn2\_6/ptodata/1/pna/US6038\_COMB.seq.\*  
108: /cgn2\_6/ptodata/1/pna/US6039\_COMB.seq.\*  
109: /cgn2\_6/ptodata/1/pna/US6040\_COMB.seq.\*  
110: /cgn2\_6/ptodata/1/pna/US6041\_COMB.seq.\*  
111: /cgn2\_6/ptodata/1/pna/US6042\_COMB.seq.\*  
112: /cgn2\_6/ptodata/1/pna/US6043\_COMB.seq.\*  
113: /cgn2\_6/ptodata/1/pna/US6044\_COMB.seq.\*  
114: /cgn2\_6/ptodata/1/pna/US6045\_COMB.seq.\*  
115: /cgn2\_6/ptodata/1/pna/US6046\_COMB.seq.\*  
116: /cgn2\_6/ptodata/1/pna/US6046\_COMB.seq.\*

117: /cgn2\_6/ptodata/1/pna/US6047\_COMB.seq.\*  
118: /cgn2\_6/ptodata/1/pna/US6048\_COMB.seq.\*  
119: /cgn2\_6/ptodata/1/pna/US6049\_COMB.seq.\*  
120: /cgn2\_6/ptodata/1/pna/US6050\_COMB.seq.\*  
121: /cgn2\_6/ptodata/1/pna/US6051\_COMB.seq.\*  
122: /cgn2\_6/ptodata/1/pna/US6052\_COMB.seq.\*  
123: /cgn2\_6/ptodata/1/pna/US6053\_COMB.seq.\*  
124: /cgn2\_6/ptodata/1/pna/US6054\_COMB.seq.\*  
125: /cgn2\_6/ptodata/1/pna/US6055\_COMB.seq.\*  
126: /cgn2\_6/ptodata/1/pna/US6056\_COMB.seq.\*  
127: /cgn2\_6/ptodata/1/pna/US6057\_COMB.seq.\*  
128: /cgn2\_6/ptodata/1/pna/US6058\_COMB.seq.\*  
129: /cgn2\_6/ptodata/1/pna/US6059\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	99.9	1253	17	US-09-155-676A-4
2	1252	99.9	1253	17	Sequence 4, Appli
3	1252	99.9	1253	124	Sequence 1553, Ap
4	1252	99.9	1253	124	Sequence 1554, Ap
5	1252	99.9	1253	124	Sequence 1555, Ap
6	1252	99.9	1253	124	Sequence 1556, Ap
7	1021.8	81.5	1200	22	US-09-471-275-1856
8	1021.8	81.5	1200	22	Sequence 1856, Ap
9	1021.8	81.5	1200	26	US-09-488-725B-6991
10	997.4	79.6	1355	33	Sequence 6991, Ap
11	997.4	79.6	1355	33	Sequence 382, App
12	663.4	52.9	3215	90	US-10-217-607-382
13	663.4	52.9	3215	90	Sequence 382, App
14	663.4	52.9	3215	90	Sequence 7088, Ap
15	661.8	52.8	2561	102	Sequence 1179, Ap
16	656	52.4	1191	97	Sequence 2059, Ap
17	648.2	51.7	1637	31	Sequence 7629, Ap
18	648.2	51.7	1637	31	Sequence 3149, Ap
19	644.4	51.4	1262	18	Sequence 8348, Ap
20	644.4	51.4	1262	18	Sequence 1616, Ap
21	644.4	51.4	1262	38	Sequence 15304, A
22	644.4	51.4	1262	38	Sequence 15304, A
23	634.8	50.7	3252	97	Sequence 15304, A
24	595	47.5	1318	21	Sequence 15304, A
25	595	47.5	1318	21	Sequence 9889, Ap
26	595	47.5	1318	39	Sequence 9889, Ap
27	580.8	46.4	2273	86	Sequence 12955, A
28	548.6	43.8	881	92	Sequence 12955, A
29	548.6	43.8	881	102	Sequence 12955, A
30	514.6	41.1	681	1	Sequence 3752, Ap
31	487.4	38.9	577	24	Sequence 19571, A
32	487	38.9	764	30	Sequence 31548, A
33	479.6	38.3	2557	1	Sequence 14062, A
34	451.8	36.1	687	49	Sequence 23046, A
35	451.8	36.1	3948	39	Sequence 38959, A
36	451.8	36.1	3948	52	Sequence 16, Appl
37	451.8	36.1	3948	58	Sequence 16, Appl
38	451.4	36.0	687	115	Sequence 1425, Ap
39	451.4	36.0	687	115	Sequence 1425, Ap
40	451.4	36.0	687	116	Sequence 1425, Ap
41	444.4	35.5	500	47	Sequence 45728, A
42	444.4	35.5	500	50	Sequence 45728, A
43	428.2	34.2	444	21	Sequence 4897, Ap
44	428.2	34.2	444	40	Sequence 4897, Ap
45	407	32.5	430	20	Sequence 15129, A

ALIGNMENTS

RESULT 1  
US-09-155-676-4  
; Sequence 4, Application US/09155676  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/155,676  
; FILING DATE: 04-JAN-1999  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IL97/00117  
; FILING DATE: 01-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 117800  
; FILING DATE: 02-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 119133  
; FILING DATE: 26-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25,618  
; REFERENCE/DOCKET NUMBER: WALLACH=21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1253 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-155-676-4  
Query Match 99.9%; Score 1252; DB 17; Length 1253;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CATTGGAGTCACGGCTGGCGGCTCTAGAAATAGTAGGATCCCGGGCTGCANGGAATTC 60  
Db 1 CATTGGAGTCACGGCTGGCGGCTCTAGAAATAGTAGGATCCCGGGCTGCANGGAATTC 60  
QY 61 GATTTCAGAGCCCAAGAGGCCCTTCTTCTGTGGTGGCGGACGTTTACAGCCGACGAC 120  
Db 61 GATTTCAGAGCCCAAGAGGCCCTTCTTCTGTGGTGGCGGACGTTTACAGCCGACGAC 120  
QY 121 CCAGCGGAGCTCAGAGGAGGCTTTTGGAGAGGCTCTCTGCCCCAGGTGGAGCGGCCCGCAA 180  
Db 121 CCAGCGGAGCTCAGAGGAGGCTTTTGGAGAGGCTCTCTGCCCCAGGTGGAGCGGCCCGCAA 180  
QY 181 GGCATCCCGCGCGCTCAGGTGGAGGCTATGTGCCCCGAACACGAGCGATGCTGCTGGTG 240  
Db 181 GGCATCCCGCGCGCTCAGGTGGAGGCTATGTGCCCCGAACACGAGCGATGCTGCTGGTG 240

```
Qy 241 CCTGTGCTCGGCTGTGAGTTCGGGACACCTGAGCCATGAAACCTGACGGTCTGTA 300
Db 241 CCTGTGCTCGGCTGTGAGTTCGGGACACCTGAGCCATGAAACCTGACGGTCTGTA 300
Qy 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATTCG 360
Db 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATTCG 360
Qy 361 GTGGGAGAACAAAGCTGAGTCCAGATGAAGAGAGAGTTTCTGTGTCATCTCCCAAGATTA 420
Db 361 GTGGGAGAACAAAGCTGAGTCCAGATGAAGAGAGAGTTTCTGTGTCATCTCCCAAGATTA 420
Qy 421 TGCCGATTCAGAAATCCATGTTGAAAGTTTGGATTCTATGAAGAAAGGAGGATAA 480
Db 421 TGCCGATTCAGAAATCCATGTTGAAAGTTTGGATTCTATGAAGAAAGGAGGATAA 480
Qy 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGAGCAGAGCCAGAGGAGTGGT 540
Db 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGAGCAGAGCCAGAGGAGTGGT 540
Qy 541 TCGGTCTGTCTTAGAGCTTCAGGAGTCCAGAGCCAGAGAGGGCTCTTCAGCACCTAG 600
Db 541 TCGGTCTGTCTTAGAGCTTCAGGAGTCCAGAGCCAGAGAGGGCTCTTCAGCACCTAG 600
Qy 601 AAGCTGGAAGGGATGAACAGCAAGTAGCTTCAGCTTACAGCAGCCCTCAAAATTTGA 660
Db 601 AAGCTGGAAGGGATGAACAGCAAGTAGCTTCAGCTTACAGCAGCCCTCAAAATTTGA 660
Qy 661 CCTGCCACAGCTCCAGAGCTTGACTGATGGAGACAGGACCATCTCGACATTCATTGG 720
Db 661 CCTGCCACAGCTCCAGAGCTTGACTGATGGAGACAGGACCATCTCGACATTCATTGG 720
Qy 721 CCATCAGGATATACAGGAGTTGTTAAATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Db 721 CCATCAGGATATACAGGAGTTGTTAAATCCACTCAGGTGCCACACCTCCCTGGATGAT 780
Qy 781 CCAAGATGAAGATACATGTCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Db 781 CCAAGATGAAGATACATGTCTGGGAACCAAGAAATAGGACCATCTATGAAGAAATTTCT 840
Qy 841 TAAAGAAAGGAAACAGAGATTGAAAGAACTCCCCAGACCGAGTTGGGGCCAACTT 900
Db 841 TAAAGAAAGGAAACAGAGATTGAAAGAACTCCCCAGACCGAGTTGGGGCCAACTT 900
Qy 901 TGATCAGCTCCAGGACAGTGCAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATA 960
Db 901 TGATCAGCTCCAGGACAGTGCAGGCTGGCTGCTCTTTTGGGCGCGCTCTGGAATA 960
Qy 961 ATGAGCGCGCTGGCAGTCCAGATCAATCACTCCAAAATGAACTGAGCTGAGCAATGAAGA 1020
Db 961 ATGAGCGCGCTGGCAGTCCAGATCAATCACTCCAAAATGAACTGAGCTGAGCAATGAAGA 1020
Qy 1021 CAGTCACATACAGAAAGCTTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080
Db 1021 CAGTCACATACAGAAAGCTTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCC 1080
Qy 1081 AAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTCTTAG 1140
Db 1081 AAGTCAACCAACCCCTATTATACCTTCCACCCAAATTTCTTATCATTTCTTCTTAG 1140
Qy 1141 GAAACAGACATACATTCATTGATTAATAAGTTTTATTTTTCGGCCCTTCGTGGCCT 1200
Db 1141 GAAACAGACATACATTCATTGATTAATAAGTTTTATTTTTCGGCCCTTCGTGGCCT 1200
Qy 1201 CGAATCAAGCTTATCGATACCGTGCACCTCGAGGGGGGGCGGTACCCACTTTT 1253
Db 1201 CGAATCAAGCTTATCGATACCGTGCACCTCGAGGGGGGGCGGTACCCACTTTT 1253
```

RESULT 2

US-09-155-676A-4

; Sequence 4, Application US/09155676A

; GENERAL INFORMATION:

```
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
; APPLICANT: BOLDIN, Mark
; APPLICANT: KOVALENKO, Andrei
; APPLICANT: METT, Igor
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
; STREET: 624 Ninth Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/155,676A
; FILING DATE: 04-JAN-1999
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 117800
; FILING DATE: 02-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 119133
; FILING DATE: 26-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: WALLACH-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; LENGTH: 1253 base pairs
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-155-676A-4
```

```
Query Match 99.9%; Score 1252; DB 17; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CATTGGAGTCAAGGAGGCTCTTCTGTGTCGGGACAGTTTACAGCGCGCAAGCAC 60
Db 1 CATTGGAGTCAAGGAGGCTCTTCTGTGTCGGGACAGTTTACAGCGCGCAAGCAC 60
Qy 61 GATTTCGAGCCACCAAGCGCCCTTCTTCTGTGTCGGGACAGTTTACAGCGCGCAAGCAC 120
Db 61 GATTTCGAGCCACCAAGCGCCCTTCTTCTGTGTCGGGACAGTTTACAGCGCGCAAGCAC 120
Qy 121 CCAGCGGAGCTGAAGGAGGCTTTTGAGAGGCTCTTCTGCCCCAGGTGGAGGGCGCCGCAA 180
Db 121 CCAGCGGAGCTGAAGGAGGCTTTTGAGAGGCTCTTCTGCCCCAGGTGGAGGGCGCCGCAA 180
Qy 181 GGCCATCCGGCCCGCTCAGGTGGAGCGCTATGTGCCCGAACAACAGAGCGATGCTGCTGTG 240
Db 181 GGCCATCCGGCCCGCTCAGGTGGAGCGCTATGTGCCCGAACAACAGAGCGATGCTGCTGTG 240
Qy 241 CCTGTGCTCGGCTGTGAGGTGCGGGAACACCTCAGCCATCGAAACCTGACGGTCTGTA 300
Db 241 CCTGTGCTCGGCTGTGAGGTGCGGGAACACCTCAGCCATCGAAACCTGACGGTCTGTA 300
```

```
QY 301 CGGGGGCTGTGGAGCTCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATTCG 360
Db |||||||
QY 301 CGGGGGCTGTGGAGCTCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATTCG 360
Db |||||||
QY 361 GTGGGAGAACAAAGCTGAGGTCAGATGAAGAGAGTTTCTGTGTCATCTCCCAAGATTA 420
Db |||||||
QY 361 GTGGGAGAACAAAGCTGAGGTCAGATGAAGAGAGTTTCTGTGTCATCTCCCAAGATTA 420
Db |||||||
QY 421 TCGCGGATCAAGAAATCCATGGTGAAGTTTGAATTCCTATGAAGAAAGGAGGATAA 480
Db |||||||
QY 421 TCGCGGATCAAGAAATCCATGGTGAAGTTTGAATTCCTATGAAGAAAGGAGGATAA 480
Db |||||||
QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 540
Db |||||||
QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 540
Db |||||||
QY 541 TCGGTCCTGTCTTAGAGCTTCAGGCTCAGAGCCAGAGAGGAGGCTCTTCAGCACCTAG 600
Db |||||||
QY 541 TCGGTCCTGTCTTAGAGCTTCAGGCTCAGAGCCAGAGAGGAGGCTCTTCAGCACCTAG 600
Db |||||||
QY 601 AAGCTGGAAGGATGAACAGCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGA 660
Db |||||||
QY 601 AAGCTGGAAGGATGAACAGCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGA 660
Db |||||||
QY 661 CTTGCCACCAAGCTCCAGAGCTTGAATGGATGGAGCAGAGCCATCTCTGACATTCATTGG 720
Db |||||||
QY 661 CTTGCCACCAAGCTCCAGAGCTTGAATGGATGGAGCAGAGCCATCTCTGACATTCATTGG 720
Db |||||||
QY 721 CCATCAGGATATACCAAGGAGTTGGTAACATCACTCAGGTGCAACCTCCCTGGATGAT 780
Db |||||||
QY 721 CCATCAGGATATACCAAGGAGTTGGTAACATCACTCAGGTGCAACCTCCCTGGATGAT 780
Db |||||||
QY 781 CCAAGATGAAGATCAATTCCTGGGAAACCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db |||||||
QY 781 CCAAGATGAAGATCAATTCCTGGGAAACCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db |||||||
QY 841 TAAAGAAAGGAAAAACAGAAAGTTGAAAAAATCTCCCGCCAGAGCCGAGTTGGGGCCAACTT 900
Db |||||||
QY 841 TAAAGAAAGGAAAAACAGAAAGTTGAAAAAATCTCCCGCCAGAGCCGAGTTGGGGCCAACTT 900
Db |||||||
QY 901 TGATCAGCTCAGAGCAAGTCAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 960
Db |||||||
QY 901 TGATCAGCTCAGAGCAAGTCAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 960
Db |||||||
QY 961 ATGAGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1020
Db |||||||
QY 961 ATGAGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1020
Db |||||||
QY 1021 CAGTCAATACAGAAAAAGCTAATCATGCTCTTACCAACTGAGCTGAGCTGAGCTGAGCTGAGCT 1080
Db |||||||
QY 1021 CAGTCAATACAGAAAAAGCTAATCATGCTCTTACCAACTGAGCTGAGCTGAGCTGAGCTGAGCT 1080
Db |||||||
QY 1081 AAGTCAACCAACCCCTATTATACCTTCCACCAAAATTTTATCATTTCTTTCTTTAG 1140
Db |||||||
QY 1081 AAGTCAACCAACCCCTATTATACCTTCCACCAAAATTTTATCATTTCTTTCTTTAG 1140
Db |||||||
QY 1141 GAAACAGACATCTCATTTGATTTGATTTAATAAGTTTATTTTTCGGCTTCTGTCGCT 1200
Db |||||||
QY 1141 GAAACAGACATCTCATTTGATTTGATTTAATAAGTTTATTTTTCGGCTTCTGTCGCT 1200
Db |||||||
QY 1201 CGAATCAAGCTTATCGATACCGTTCAGCTCGAGGGGGGGCGGACCACTTTT 1253
Db |||||||
QY 1201 CGAATCAAGCTTATCGATACCGTTCAGCTCGAGGGGGGGCGGACCACTTTT 1253
Db |||||||
```

```
RESULT 3
US-60-545-213-1553
; Sequence 1553, Application US/60545213
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William Martin
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug
; TARGET GENES: Target Genes
```

```
; FILE REFERENCE: AM101083 (031896-042099)
; CURRENT APPLICATION NUMBER: US/60/545,213
; CURRENT FILING DATE: 2004-02-18
; NUMBER OF SEQ ID NOS: 303284
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1553
; LENGTH: 1253
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: n is a, c, g, or t
US-60-545-213-1553
```

```
Query Match 99,9%; Score 1252; DB 124; Length 1253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATTGGAGTCAACGCGTGGCGGCTCTAGAAATAGTGGATCCCGGGCTGCAGGAATTC 60
Db |||||||
QY 61 GATTGAGCCCAACGAGCCCTTCTTCTGTGTGCGGCGACGTTTACAGCCGCAAGCAC 120
Db |||||||
QY 61 GATTGAGCCCAACGAGCCCTTCTTCTGTGTGCGGCGACGTTTACAGCCGCAAGCAC 120
Db |||||||
QY 121 CCAGCGGAGCTGAAGAGAGGCTTTTGAGAGGCTCTTCCCGCAGGTGGAGGCGGCGCA 180
Db |||||||
QY 121 CCAGCGGAGCTGAAGAGAGGCTTTTGAGAGGCTCTTCCCGCAGGTGGAGGCGGCGCA 180
Db |||||||
QY 181 GGCCATCCGCGCTCAGGTGGAGGCTATGTGCCCGAACAAGAGGATCTGCTGTG 240
Db |||||||
QY 181 GGCCATCCGCGCTCAGGTGGAGGCTATGTGCCCGAACAAGAGGATCTGCTGTG 240
Db |||||||
QY 241 CTTGTGCTGCGCTGAGGTGCGGAACTCTGAGCCATGAAACCTGACGCTGCTGTA 300
Db |||||||
QY 241 CTTGTGCTGCGCTGAGGTGCGGAACTCTGAGCCATGAAACCTGACGCTGCTGTA 300
Db |||||||
QY 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATTCG 360
Db |||||||
QY 301 CGGGGGCTGTGGAGCATCTGGCCAGCCAGAGCACAAGAAAGCAACCAAAATTCG 360
Db |||||||
QY 361 GTGGAGAACAAAGCTGAGGTCAGATGAAGAGAGTTTCTGTGTCATCTCCCAAGATTA 420
Db |||||||
QY 361 GTGGAGAACAAAGCTGAGGTCAGATGAAGAGAGTTTCTGTGTCATCTCCCAAGATTA 420
Db |||||||
QY 421 TCGCGGATCAAGAAATCCATGGTGAAGTTTGAATTCCTATGAAGAAAGGAGGATAA 480
Db |||||||
QY 421 TCGCGGATCAAGAAATCCATGGTGAAGTTTGAATTCCTATGAAGAAAGGAGGATAA 480
Db |||||||
QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 540
Db |||||||
QY 481 AGTGATCAAGAGATGGCAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTGGT 540
Db |||||||
QY 541 TCGGTCCTGTCTTAGAGCTTCAGGCTCAGAGCCAGAGAGGCTCTTCAGCACCTAG 600
Db |||||||
QY 541 TCGGTCCTGTCTTAGAGCTTCAGGCTCAGAGCCAGAGAGGCTCTTCAGCACCTAG 600
Db |||||||
QY 601 AAGCTGGAAGGATGAACAGCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGA 660
Db |||||||
QY 601 AAGCTGGAAGGATGAACAGCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGA 660
Db |||||||
QY 661 CTTGCCACCAAGCTCCAGAGCTTGAATGGATGGAGCAGAGCCATCTCTGACATTCATTGG 720
Db |||||||
QY 661 CTTGCCACCAAGCTCCAGAGCTTGAATGGATGGAGCAGAGCCATCTCTGACATTCATTGG 720
Db |||||||
QY 721 CCATCAGGATATACCAAGGAGTTGGTAACATCACTCAGGTGCAACCTCCCTGGATGAT 780
Db |||||||
QY 721 CCATCAGGATATACCAAGGAGTTGGTAACATCACTCAGGTGCAACCTCCCTGGATGAT 780
Db |||||||
QY 781 CCAAGATGAAGATCAATTCCTGGGAAACCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db |||||||
QY 781 CCAAGATGAAGATCAATTCCTGGGAAACCAAGAAATAGGACCATCTATGAAGATTTCT 840
Db |||||||
```





Db 121 CCAGCGGAGCTGAAGAGGCTTTTGTAGAGGCTCTCTGCCCCAGGTGGAGGCGGCCGCGAA 180  
Qy 181 GCCCATCCGCGCGCTCAGAGTGGAGCTATGTGCCCCGAAACACGAGGAGTCTGCTGGTG 240  
Db 181 GCCCATCCGCGCGCTCAGAGTGGAGGCTATGTGCCCCGAAACACGAGGAGTCTGCTGGTG 240  
Qy 241 CCTGTGCTGCGGCTGTGAGGTGCGGGAAACACTGAGCCATGGAACCTGACGGTCTGTGA 300  
Db 241 CCTGTGCTGCGGCTGTGAGGTGCGGGAAACACTGAGCCATGGAACCTGACGGTCTGTGA 300  
Qy 301 CGGGGGGCTGTGAGGATCTGGCCAGCCAGAGCACAAGAAAGCAACCAACCAATCTG 360  
Db 301 CGGGGGGCTGTGAGGATCTGGCCAGCCAGAGCACAAGAAAGCAACCAACCAATCTG 360  
Qy 361 GTGGGAGAACAAAGCTGAGGTCCAGATGAAAGAAAGTTTCTGCTCACTCCGCCAGGATTA 420  
Db 361 GTGGGAGAACAAAGCTGAGGTCCAGATGAAAGAAAGTTTCTGCTCACTCCGCCAGGATTA 420  
Qy 421 TGGCGGATTCAAGAAATCCATGGTGAAGTTTGGATTCTTATGAAGAAAGGAGGATTA 480  
Db 421 TGGCGGATTCAAGAAATCCATGGTGAAGTTTGGATTCTTATGAAGAAAGGAGGATTA 480  
Qy 481 AGTCAATCAAGGAGATGGCAGCTCAGATCCGTTAGTGGAGCAGAGCCGACAGGAGTGGT 540  
Db 481 AGTCAATCAAGGAGATGGCAGCTCAGATCCGTTAGTGGAGCAGAGCCGACAGGAGTGGT 540  
Qy 541 TCGGCTGTCTTAGAGCTCAGGAGTGCAGAGCCAGAGAGGGCTTTCAGCACCTAG 600  
Db 541 TCGGCTGTCTTAGAGCTCAGGAGTGCAGAGCCAGAGAGGGCTTTCAGCACCTAG 600  
Qy 601 AAGCTGAAAGGGATGAACAGCAAGTAGTTCAGCTTACAGAGCCCTCAAAATTTGGA 660  
Db 601 AAGCTGAAAGGGATGAACAGCAAGTAGTTCAGCTTACAGAGCCCTCAAAATTTGGA 660  
Qy 661 CCTGCCACCACTCCAGAGCTTGATGGATGGAGACAGGACCATCTCTGACATTCATGG 720  
Db 661 CCTGCCACCACTCCAGAGCTTGATGGATGGAGACAGGACCATCTCTGACATTCATGG 720  
Qy 721 CCATCAGATATACAGAGTTGGTAAATCCATCCACTCAGGTGCCACACCTCCCTGGATGAT 780  
Db 721 CCATCAGATATACAGAGTTGGTAAATCCATCCACTCAGGTGCCACACCTCCCTGGATGAT 780  
Qy 781 CCAAGATGAAGAAATACATTCGTGGGAACCAAGAAATAGGACCATCTTATGAAGAAATTC 840  
Db 781 CCAAGATGAAGAAATACATTCGTGGGAACCAAGAAATAGGACCATCTTATGAAGAAATTC 840  
Qy 841 TAAAGAAAGGAAAAACAGAGTTGAAAAAATCTCCGCCAGACCGAGTTGGGGCCAACTT 900  
Db 841 TAAAGAAAGGAAAAACAGAGTTGAAAAAATCTCCGCCAGACCGAGTTGGGGCCAACTT 900  
Qy 901 TGATCAGAGTCCAGGACCACTGAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGGAATA 960  
Db 901 TGATCAGAGTCCAGGACCACTGAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGGAATA 960  
Qy 961 ATGGACGCGCTGCGAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGAAG 1020  
Db 961 ATGGACGCGCTGCGAGTCCAGACATCAACTCCAAAACTGAAGCTGCAGCAATGAAGAAG 1020  
Qy 1021 CAGTCAATACAGAAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCC 1080  
Db 1021 CAGTCAATACAGAAAAAAGCTAATCATGCTCTCTACCAACTACCATGAGGCTAAAAGCC 1080  
Qy 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCCAAATCTTTATCATGTTCTTTCTTAG 1140  
Db 1081 AAAGTCAACAAACCCCTATTATACCTTCCACCCAAATCTTTATCATGTTCTTTCTTAG 1140  
Qy 1141 GAAACAGACATCACTCATTTGATTTAATAAAGTTTATTTTTCGGCCCTTCGTGGCCT 1200  
Db 1141 GAAACAGACATCACTCATTTGATTTAATAAAGTTTATTTTTCGGCCCTTCGTGGCCT 1200  
Qy 1201 CGAATCAAGCTTATCGATACCGTCACTCGAGGGGGGGCGGTATCCCACTTTT 1253  
Db 1201 CGAATCAAGCTTATCGATACCGTCACTCGAGGGGGGGCGGTATCCCACTTTT 1253

RESULT 7

US-09-471-275-1856/c  
; Sequence 1856, Application US/09471275  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: Novel Contigs Obtained  
; TITLE OF INVENTION: From Various Libraries  
; FILE REFERENCE: 782  
; CURRENT APPLICATION NUMBER: US/09/471,275  
; EARLIER FILING DATE: 1999-12-23  
; EARLIER APPLICATION NUMBER: US 09/235,076  
; EARLIER FILING DATE: 1999-01-20  
; EARLIER APPLICATION NUMBER: US 09/234,611  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER APPLICATION NUMBER: US 09/240,371  
; EARLIER FILING DATE: 1999-01-29  
; EARLIER APPLICATION NUMBER: US 09/277,227  
; EARLIER FILING DATE: 1999-03-25  
; EARLIER APPLICATION NUMBER: US 09/271,490  
; EARLIER FILING DATE: 1999-03-18  
; EARLIER APPLICATION NUMBER: US 09/293,972  
; EARLIER FILING DATE: 1999-04-15  
; EARLIER APPLICATION NUMBER: US 09/274,861  
; EARLIER FILING DATE: 1999-03-23  
; EARLIER APPLICATION NUMBER: US 60/125,453  
; EARLIER FILING DATE: 1999-03-19  
; EARLIER APPLICATION NUMBER: US 60/126,605  
; EARLIER FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: US 09/306,350  
; EARLIER FILING DATE: 1999-05-07  
; EARLIER APPLICATION NUMBER: US 09/399,720  
; EARLIER FILING DATE: 1999-09-21  
; EARLIER APPLICATION NUMBER: US 09/404,284  
; EARLIER FILING DATE: 1999-09-21  
; EARLIER APPLICATION NUMBER:  
; NUMBER OF SEQ ID NOS: 10451  
; SOFTWARE: pt CT\_genes Version 1.0  
; SEQ ID NO 1856  
; LENGTH: 1200  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(1201)  
; OTHER INFORMATION: n = a,t,c or g  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1135)...(127)  
; OTHER INFORMATION: similar to gi4455233 in the genept database release 114,  
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters

US-09-471-275-1856

Query Match 81.5%; Score 1021.8; DB 22; Length 1200;  
Best Local Similarity 98.7%; Pred. No. 4.5e-262;  
Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;  
Qy 81 CCTTCTTCTGTGTGTC-GGCAGCTTTACAGCCCAAGCACCCAGCGGCTGAAGGAG 139  
Db 1115 CCTTCTTCTGTGTGCGGCGCAGCTTTACAGCCCAAGCA-CCAGCGCAGCTGAAGGAG 1057  
Qy 140 GCCTTTGAGAGGCTCTCTGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAG 199  
Db 1056 GCCTTTGAGAGGCTCTCTGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAG 997  
Qy 200 GTGAGGCGCTATGTGCCCAACACGAGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGAG 259  
Db 996 GTGAGGCGCTATGTGCCCAACACGAGCGATGCTGTGCTGTGCTGTGCTGTGCTGTGAG 937  
Qy 260 GTGCGGGAACACCTGAGGCCATGGAACCTCAAGTGTGTACCGGGGGGCTGCTGGAGCAT 319  
Db 1115 GTGCGGGAACACCTGAGGCCATGGAACCTCAAGTGTGTACCGGGGGGCTGCTGGAGCAT 319



```
Db 936 GTGCGGAAACACCTGAGCCATGAAACCTGACGGTCTGTACGGGGGGCTGCTGGAGCAT 877
QY 320 CTGGCCAGCCCGCAGAGCACAGAAAGCAACCAAAATCTGTGGGAGAACAAAGCTGAG 379
Db 876 CTGGCCAGCCCGCAGAGCACAGAAAGCAACCAAAATCTGTGGGAGAACAAAGCTGAG 817
QY 380 GTCCAGATGAAGAGAGAGTTTCTGTGTCACCTCCCGAGATTATGCGGATTCGAAGAAATCC 439
Db 816 GTCCAGATGAAGAGAGAGTTTCTGTGTCACCTCCCGAGATTATGCGGATTCGAAGAAATCC 757
QY 440 ATGGTCAAAAGTTTGAATTCCTATGAAGAAAGAGGATTAAGTGAATCAAGAGATGGCA 499
Db 756 ATGGTCAAAAGTTTGAATTCCTATGAAGAAAGAGGATTAAGTGAATCAAGAGATGGCA 697
QY 500 GCTCAGATCCGTGAGTGAGCAGAGCCGACAGGAGGTGTTCTGGTCTGTCTAGAGCCT 559
Db 696 GCTCAGATCCGTGAGTGAGCAGAGCCGACAGGAGGTGTTCTGGTCTGTCTAGAGCCT 637
QY 560 CAGGAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGGGATGAAC 619
Db 636 CAGGAGTGCCAGACCCAGAGAGGGCTCTTCAGCACCTAGAGCTGGAAGGGATGAAC 577
QY 620 AGCCAAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTTGGACCTGACACAGCTCCAGAG 679
Db 576 AGCCAAAGTAGCTTCAGCTTACAGAGCCCTCAAAATTTGGACCTGACACAGCTCCAGAG 517
QY 680 CTTGACTGGATGAGACAGGACCATCTCTGACATTCATTTGGCCATCAGGATATACAGGA 739
Db 516 CTTGACTGGATGAGACAGGACCATCTCTGACATTCATTTGGCCATCAGGATATACAGGA 457
QY 740 GTTGTAAACATCCACTCAGTGCCACACCTCCCTGGATGATCCAAGATGAAGAATACATT 799
Db 456 GTTGTAAACATCCACTCAGTGCCACACCTCCCTGGATGATCCAAGATGAAGAATACATT 397
QY 800 GCTGGGAACCAAGAAATAGGACATCTCTATGAAGAAATTTCTTAAAGAAAGGAAACAG 859
Db 396 GCTGGGAACCAAGAAATAGGACATCTCTATGAAGAAATTTCTTAAAGAAAGGAAACAG 337
QY 860 AAGTTGAAAAAATCTCCCCAGACGAGTTGGGGCCAACTTTGATCAGAGCTCCAGGACC 919
Db 336 AAGTTGAAAAAATCTCCCCAGACGAGTTGGGGCCAACTTTGATCAGAGCTCCAGGACC 277
QY 920 AGTCAGAGTGCTGCTCTCTTTTGGCGCGCTCTGGAATAATGGACGCGCTGGCAGTC 979
Db 276 AGTCAGAGTGCTGCTGCTCTCTTTTGGCGCGCTCTGGAATAATGGACGCGCTGGCAGTC 218
QY 980 CAGACATCAACTCCAAACTGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAA 1039
Db 217 CAGACATCAA--TCAAAACTGAAGCTGAGCAATGAAGAGCAGTCAATACAGAAAAAA 159
QY 1040 GCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGCCAAAGTCAACCAAAACCCCTA 1099
Db 158 GCTAATCATGCTCTCTACCACTACCATGAGGCTAAAGC--AAAGTCAACAAACCCCTA 101
QY 1100 TTATACCTTCCACCCAAATCTTTATCATTTCTTTTAGGAAACAGACATACCTCATTC 1159
Db 100 TTATACCTTCCA--CCAAATCTTTATCATTTCTTTTAGGAAACAGACATACCTCATTC 42
QY 1160 ATTTGATTAATAAGTTTATTTTTC 1186
Db 41 ATTTGATTAATAAGTTTATTTTTC 15
```

```
RESULT 8
US-09-488-725B-6991/c
; Sequence 6991, Application US/09488725B
; GENERAL INFORMATION:
; APPLICANT: Yuanhua T. Tang
; APPLICANT: John Tillinghast
; APPLICANT: Ankura Sinku
; APPLICANT: Chenghua Liu
; APPLICANT: Radoje T. Drmanac
; TITLE OF INVENTION: Novel Contigs Obtained
```

```
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 784
; CURRENT APPLICATION NUMBER: US/09/488,725B
; CURRENT FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: US 09/034,341
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 09/045,400
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: US 09/131,598
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: US 09/170,294
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 09/179,473
; PRIOR FILING DATE: 1998-10-27
; PRIOR APPLICATION NUMBER: US 09/181,430
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: US 09/235,076
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: US 09/234,611
; PRIOR FILING DATE: 1999-01-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10289
; SOFTWARE: pt_CT_genes Version 1.01
; SEQ ID NO 6991
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1201)
; OTHER INFORMATION: n = a,t,c or g
; NAME/KEY: misc feature
; LOCATION: (1135)...(127)
; OTHER INFORMATION: similar to gi4455233 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-488-725B-6991
```

```
Query Match 81.5%; Score 1021.8; DB 22; Length 1200;
Best Local Similarity 98.7%; Pred. No. 4.5e-262;
Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;

QY 81 CCTCTCTTCTGTGTGTCG--GGCAGCTTTACAGCGCGCAAGCACCCAGCGCGAGCTGAAGGAG 139
Db 1115 CCTCTCTTCTGTGTGTCGCGGCACGTTTACAGCGCGAGCA--CCAGCGCGAGCTGAAGGAG 1057
QY 140 GCTTTTGAGAGGCTCCTGCCCGCAGGTGGAGGGCGCCGCAAGGCCATCCGCGCGCTCAG 199
Db 1056 GCTTTGGAGAGGCTCCTGCCCGCAGGTGGAGGGCGCCGCAAGGCCATCCGCGCGCTCAG 997
QY 200 GTGAGCGCTATGTGCCCCGAAACACAGAGCGATGCTGCTGGTGCCTGTGCTGGGCTGTGAG 259
Db 956 GTGAGCGCTATGTGCCCCGAAACACAGAGCGATGCTGCTGGTGCCTGTGCTGGGCTGTGAG 937
QY 260 GTGCGGGAAACCTTGAGCCATGGAACCTGACGGTGTGTACGGGGGGCTGCTGGAGCAT 319
Db 936 GTGCGGGAAACCTTGAGCCATGGAACCTGACGGTGTGTACGGGGGGCTGCTGGAGCAT 877
QY 320 CTGGCCAGCCCGCAGAGCACAGAAAGCAACCAAAATCTGTGGGAGAACAAAGCTGAG 379
Db 876 CTGGCCAGCCCGCAGAGCACAGAAAGCAACCAAAATCTGTGGGAGAACAAAGCTGAG 817
QY 380 GTCCAGATGAAGAGAGAGTTTCTGTGTCACCTCCCGAGATTATGCGGATTCGAAGAAATCC 439
Db 816 GTCCAGATGAAGAGAGAGTTTCTGTGTCACCTCCCGAGATTATGCGGATTCGAAGAAATCC 757
QY 440 ATGGTGAAGGTTTGGATTCTCTATGAAGAAAGAGGATTAAGTGAATCAAGAGATGGCA 499
Db 756 ATGGTGAAGGTTTGGATTCTCTATGAAGAAAGAGGATTAAGTGAATCAAGAGATGGCA 697
```



```
Qy 500 GCTCAGATCCGTGAGGTGGAGCAGACGAGCGCAGAGAGGTGGTTCGGTCTGTCTTAGAGCCT 559
Db 696 GCTCAGATCCGTGAGGTGGAGCAGACGAGCGCAGAGAGGTGGTTCGGTCTGTCTTAGAGCCT 637
Qy 560 CAGCAGTGGCCAGACCCAGAGAGGGCTCTTTCAGCAGCTAGAGCTGGAAGGGATGAAC 619
Db 636 CAGCAGTGGCCAGACCCAGAGAGGGCTCTTTCAGCAGCTAGAGCTGGAAGGGATGAAC 577
Qy 620 AGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACCAAGCTCCAGAG 679
Db 576 AGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACCAAGCTCCAGAG 517
Qy 680 CTTGACTGGATGGAGACAGACCATCTCTGACATTCATTGGGCCATCAGGATATACACGGA 739
Db 516 CTTGACTGGATGGAGACAGACCATCTCTGACATTCATTGGGCCATCAGGATATACACGGA 457
Qy 740 GTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATT 799
Db 456 GTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATT 397
Qy 800 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAAAACAG 859
Db 396 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAAAACAG 337
Qy 860 AAGTTGAAAAAATCTCCCCAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 919
Db 336 AAGTTGAAAAAATCTCCCCAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 277
Qy 920 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATTAATGGAGCGCGCTGGCAGTC 979
Db 276 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATTAATGGAGCGCGCTGGCAGTC 218
Qy 980 CAGACATCAACTCCAAAATGAAGCTGCAGCAATGAAGAGCAGTCACATACAGAAAAA 1039
Db 217 CAGACATCAA-TTCAAACTGAGCTGCAGCAATGAAGAGCAGTCACATACAGAAAAA 159
Qy 1040 GCTAATCATGCTCTTACCAACTACCATGAGGCTAAAGGCCAAAGTCAACAAACCCCTA 1099
Db 158 GCTAATCATGCTCTTACCAACTACCATGAGGCTAAAGGC--AAAGTCAACAAACCCCTA 101
Qy 1100 TTATACCTCCACCAAAATCTTTATCATCTGTCTTTCTTAGGAAACAGACATCTCATTC 1159
Db 100 TTATACCTTCCA-CCAAATCTTTATCATCTGTCTTTCTTAGGAAACAGACATCTCATTC 42
Qy 1160 ATTTGATTAAATAAGTTTATTTTC 1186
Db 41 ATTTGATTAAATAAGTTTATTTTC 15
```

## RESULT 9

```
US-09-552-317-6991/c
; Sequence 6991, Application US/09552317
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sanku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids
; TITLE OF INVENTION: and Polypeptides
; FILE REFERENCE: 784CIP
; CURRENT APPLICATION NUMBER: US/09/552,317
; CURRENT FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 10289
; SOFTWARE: pt CT_genes Version 1.01
; SEQ ID NO 6991
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; LOCATION: (1)...(1201)
; OTHER INFORMATION: n = a,t,c or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1135)...(127)
; OTHER INFORMATION: similar to gi4455233 in the genepept database release 114,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-552-317-6991
```

Query Match 81.5%; Score 1021.8; DB 26; Length 1200;

Best Local Similarity 98.7%; Pred. No. 4.5e-262;

Matches 1093; Conservative 0; Mismatches 7; Indels 7; Gaps 6;

```
Qy 81 CTTCTTCTGTGTGCG-GGCAGCGTTTACAGCCGCAAGCACCCAGCGGAGCTCAAGGAG 139
Db 1115 CTTCTTCTGTGTGCGGGCAGCGTTTACAGCCGCAAGCA-CCAGCGGAGCTCAAGGAG 1057
Qy 140 GCTTTTGGAGGGCTCCTGCCCCCAGGTGGAGCGCGCCGCAAGCCATCCGCGCCGCTCAG 199
Db 1056 GCTTTTGGAGGGCTCCTGCCCCCAGGTGGAGCGCGCCGCAAGCCATCCGCGCCGCTCAG 997
Qy 200 GTGGAGCGCTATGTGCCCGCAACACAGCGATGTCTGTGTGCTGTGCTGTGCTGTGAG 259
Db 996 GTGGAGCGCTATGTGCCCGCAACACAGCGATGTCTGTGTGCTGTGCTGTGCTGTGAG 937
Qy 260 GTGGGGAACACCTGAGCCATGGAACCTGACGGTGTCTGACGGGGGGCTGTGAGAGCAT 319
Db 936 GTGGGGAACACCTGAGCCATGGAACCTGACGGTGTCTGACGGGGGGCTGTGAGAGCAT 877
Qy 320 CTGGCCAGCCAGAGCAGCAAGAAAGCAACCAAAATTTCTGGTGGGAGAACAAAGCTGAG 379
Db 876 CTGGCCAGCCAGAGCAGCAAGAAAGCAACCAAAATTTCTGGTGGGAGAACAAAGCTGAG 817
Qy 380 GTCCAGATGAAGAGAGAGTTTCTGGTCACTCCCGAGATTTATGCGCATTCAGAAATCC 439
Db 816 GTCCAGATGAAGAGAGAGTTTCTGGTCACTCCCGAGATTTATGCGCATTCAGAAATCC 757
Qy 440 ATGTTGAAAGGTTTGGATTCCTATGAAGAAAGAGAGATTAAGTATCAAGGAGATGCA 499
Db 756 ATGTTGAAAGGTTTGGATTCCTATGAAGAAAGAGAGATTAAGTATCAAGGAGATGCA 697
Qy 500 GCTCAGATCCGTGAGGTGGAGCAGACCGCAGAGAGGTGGTTCGGTCTGTCTTAGAGCCT 559
Db 696 GCTCAGATCCGTGAGGTGGAGCAGACCGCAGAGAGGTGGTTCGGTCTGTCTTAGAGCCT 637
Qy 560 CAGCAGTGGCCAGACCCAGAGAGGGCTCTTTCAGCAGCTAGAGCTGGAAGGGATGAAC 619
Db 636 CAGCAGTGGCCAGACCCAGAGAGGGCTCTTTCAGCAGCTAGAGCTGGAAGGGATGAAC 577
Qy 620 AGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACCAAGCTCCAGAG 679
Db 576 AGCCAAGTAGCTTCCAGCTTACAGCAGCCCTCAAAATTTGGACCTGCCACCAAGCTCCAGAG 517
Qy 680 CTTGACTGGATGGAGACAGACCATCTCTGACATTCATTGGGCCATCAGGATATACACGGA 739
Db 516 CTTGACTGGATGGAGACAGACCATCTCTGACATTCATTGGGCCATCAGGATATACACGGA 457
Qy 740 GTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATT 799
Db 456 GTTGGTAACATCCACTCAGGTGCCACACCTCCCTGGATGATCCAAAGATGAAGATACATT 397
Qy 800 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAAAACAG 859
Db 396 GCTGGGAACCAAGAAATAGGACCATCTCTATGAAGATTTCTTAAAGAAAGGAAAAACAG 337
Qy 860 AAGTTGAAAAAATCTCCCCAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 919
Db 336 AAGTTGAAAAAATCTCCCCAGACCGAGTTGGGCGCAACTTTTGATCAAGCTCCAGGACC 277
Qy 920 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATTAATGGAGCGCGCTGGCAGTC 979
Db 276 AGTCAGGCTGGCTGCCCTCTTTTGGCGCGCTCTGGAATTAATGGAGCGCGCTGGCAGTC 218
```





```
QY 998 CTGAAGCTGAGCAATGAAGAAGCAGTCACATACAGAAAAAGCTAATCATGCTCTCTAC 1057
Db 1223 CTGAAGCTGAGCAATGAAGAAGCAGTCACATACAGAAAAAGCTAATCATGCTCTCTAC 1282
QY 1058 CAACTACCATGAGGCTAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAA 1117
Db 1283 CAACTACCATGAGGCTAAAGCC--AAAGTCAACCAAAACCCCTATTATACCTTCCA-CCAAA 1339
QY 1118 TTCTTTATCATGCTCTTCTTAGGAAACAGACATCTCATTTGATTTAATAAGTT 1177
Db 1340 TTCTTTATCATGCTCTTCTTAGGAAACAGACATCTCATTTGATTTAATAAGTT 1399
QY 1178 TTATTTTTC 1186
Db 1400 TTATTTTTC 1408

RESULT 13
US-60-213-362-1179/c
; Sequence 1179, Application US/60213362
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
; TITLE OF INVENTION: Identified Thereby
; FILE REFERENCE: GX-0016 P
; CURRENT APPLICATION NUMBER: US/60/213,362
; CURRENT FILING DATE: 2000-06-21
; NUMBER OF SEQ ID NOS: 8429
; SOFTWARE: PERL Program
; SEQ ID NO 1179
; LENGTH: 3215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1000026.9
; NAME/KEY: unsure
; LOCATION: 1011-1042
; OTHER INFORMATION: a, t, c, g, or other
US-60-213-362-1179

Query Match 52.9%; Score 663.4; DB 90; Length 3215;
Best Local Similarity 95.1%; Pred. No. 7.2e-166;
Matches 750; Conservative 0; Mismatches 6; Indels 33; Gaps 5;

QY 426 GATTCAGAAAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 485
Db 2587 GATTCAGAAAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 2528
QY 486 TCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGCGCAGAGGAGTGGTTCGGT 545
Db 2527 TCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGCGCAGAGGAGTGGTTCGGT 2468
QY 546 CTGCTCTTAGAG-----CCTCAGGCGAGTGCACAGCCCA 577
Db 2467 CTGCTCTTAGAGTTGGTTTCCCTCGGAGGATCCAGACACCTCAGGCGAGTGCACAGCCCA 2408
QY 578 GAAGAGGGCTCTTCAGCAGCTAGAGCTGGAAGGAGTGAACAGCAAGTAGCTTCCAGC 637
Db 2407 GAAGAGGGCTCTTCAGCAGCTAGAGCTGGAAGGAGTGAACAGCAAGTAGCTTCCAGC 2348
QY 638 TTACAGCAGCCCTCAAAATTTGACCTGCCACAGCTCCAGAGCTTGACTGGATGAGACA 697
Db 2347 TTACAGCAGCCCTCAAAATTTGACCTGCCACAGCTCCAGAGCTTGACTGGATGAGACA 2288
QY 698 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTGGTAAACATCCACTCA 757
Db 2287 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTGGTAAACATCCACTCA 2228
QY 758 GGTGCCACACCTCCCTCGGATGATCCAAGATGAAGATAATGCTGGGAAACCAAGAAATA 817
```

```
Db 2227 GGTGCCACACCTCCCTCGGATGATCCAAGATGAAGATAATGCTGGGAAACCAAGAAATA 2168
QY 818 GGACCATCTCTATGAAGAAATTTCTTAAAGAAAAAGAAAAACAGAGTTGAAAAAACTCCCC 877
Db 2167 GGACCATCTCTATGAAGAAATTTCTTAAAGAAAAAGAAAAACAGAGTTGAAAAAACTCCCC 2108
QY 878 CCAGACCGGTTGGGGCCAACTTTTGATCACAGTCCAGGACCAAGTCCAGGCTGGCTGCC 937
Db 2107 CCAGACCGGTTGGGGCCAACTTTTGATCACAGTCCAGGACCAAGTCCAGGCTGGCTGCC 2048
QY 938 TCTTTTGGGGCCGCTCTGGAATAATGAGCGCGCTGGCAGTCCAGACATCAACTCCAAAA 997
Db 2047 TCTTTT-GGCCCGCTCTGGAATAATGAGCGCGCTGGCAGTCCAGACATCAA-TTCAAAA 1990
QY 998 CTGAAGCTGACAGCAATGAAGAAGCAGTCACATACAGAAAAAGCTAAATCATGCTCTCTAC 1057
Db 1989 CTGAAGCTGACAGCAATGAAGAAGCAGTCACATACAGAAAAAGCTAAATCATGCTCTCTAC 1930
QY 1058 CAACTACCATGAGGCTAAAGCCAAAGTCAACCAAAACCCCTATTATACCTTCCACCCAAA 1117
Db 1929 CAACTACCATGAGGCTAAAGC--AAAGTCAACCAAAACCCCTATTATACCTTCCA-CCAAA 1873
QY 1118 TTCTTTATCATGCTCTTCTTAGGAAACAGACATCTCATTTGATTTAATAAGTT 1177
Db 1872 TTCTTTATCATGCTCTTCTTAGGAAACAGACATCTCATTTGATTTAATAAGTT 1813
QY 1178 TTATTTTTC 1186
Db 1812 TTATTTTTC 1804

RESULT 14
US-60-278-561-2059/c
; Sequence 2059, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 2059
; LENGTH: 3216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1000026.9
; NAME/KEY: unsure
; LOCATION: 1011-1042
; OTHER INFORMATION: a, t, c, g, or other
US-60-278-561-2059

Query Match 52.9%; Score 663.4; DB 97; Length 3216;
Best Local Similarity 95.1%; Pred. No. 7.2e-166;
Matches 750; Conservative 0; Mismatches 6; Indels 33; Gaps 5;

QY 426 GATTCAGAAAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 485
Db 2587 GATTCAGAAAAATCCATGGTGAAGGTTTGGATTCTCTATGAAGAAAGGAGGATAAAGTGA 2528
QY 486 TCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGCGCAGAGGAGTGGTTCGGT 545
Db 2527 TCAAGGAGATGCGAGCTCAGATCCGTGAGGTGGAGCAGCGCAGAGGAGTGGTTCGGT 2468
QY 546 CTGCTCTTAGAG-----CCTCAGGCGAGTGCACAGCCCA 577
```

```
Db 2467 CTGCTTAGAGTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGAGTGGCCAGACCCA 2408
Qy 578 GAAGAGGGCTCTTCAGCACCTAGAGCTGGAAAGGGATGAACAGCCAAAGTAGCTTCAGC 637
Db 2407 GAAGAGGGCTCTTCAGCACCTAGAGCTGGAAAGGGATGAACAGCCAAAGTAGCTTCAGC 2348
Qy 638 TTACAGCAGCCCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGACTGGATGGAGACA 697
Db 2347 TTACAGCAGCCCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGACTGGATGGAGACA 2288
Qy 698 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTTGGTAAACATCACTCA 757
Db 2287 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTTGGTAAACATCACTCA 2228
Qy 758 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCTGGGAAACCAAGAAATA 817
Db 2227 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCTGGGAAACCAAGAAATA 2168
Qy 818 GGACCATCTCTGATGAAGATTTCTTAAAGAAAGGAAACCAAGAGTTGAAAGAACTCCCC 877
Db 2167 GGACCATCTCTGATGAAGATTTCTTAAAGAAAGGAAACCAAGAGTTGAAAGAACTCCCC 2108
Qy 878 CCAGACCGAGTTGGGGCCAACTTTGATCAGAGTCCAGGACCAAGTGGAGGCTGGCTGCC 937
Db 2107 CCAGACCGAGTTGGGGCCAACTTTGATCAGAGTCCAGGACCAAGTGGAGGCTGGCTGCC 2048
Qy 938 TCTTTTGGCGCGCTCTGGAATATGGAACGCGCTGGCAGTCCAGACATCAACTCCAAAA 997
Db 2047 TCTTTT-GGCGCGCTCTGGAATATGGAACGCGCTGGCAGTCCAGACATCAA-TTCAAAA 1990
Qy 998 CTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1057
Db 1989 CTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1930
Qy 1058 CAATACCATCAGCTTAAGCCAAAGTCAACCAACCCCTATTATACCTTCCACCCAAA 1117
Db 1929 CAATACCATCAGCTTAAGCCAAAGTCAACCAACCCCTATTATACCTTCCCA-CCAAA 1873
Qy 1118 TCTTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCAATTCATTGATTTAATAAGTT 1177
Db 1872 TCTTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCAATTCATTGATTTAATAAGTT 1813
Qy 1178 TTATTTTC 1186
Db 1812 TTATTTTC 1804
```

```
RESULT 15
US-60-324-185-7629/c
; Sequence 7629, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
; TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
; CURRENT FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 35862
; SOFTWARE: PERL Program
; SEQ ID NO 7629
; LENGTH: 2561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1301765.9
US-60-324-185-7629
```

Query Match 52.8%; Score 661.8; DB 102; Length 2561;  
Best Local Similarity 94.9%; Pred. No. 1.7e-165;

```
Matches 749; Conservative 0; Mismatches 7; Indels 33; Gaps 5;
Qy 426 GATTCAAGAAATCCATGGTGAAGGTTTGGATTTCTTATGAAGAAAAAGAGGATAAAGTGA 485
Db 1933 GATTCAAGAAATCCATGGTGAAGGTTTGGATTTCTTATGAAGAAAAAGAGGATAAAGTGA 1874
Qy 486 TCAAGAGGATGGAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTTGGTTCGGT 545
Db 1873 TCAAGAGGATGGAGCTCAGATCCGTGAGTGGAGCAGAGCCGACAGAGGTTGGTTCGGT 1814
Qy 546 CTGTCTTTAGAG-----CCTCAGGCAGTGGCCAGACCCA 577
Db 1813 CTGTCTTTAGAGTTGGTTTCCCTCGGAGGATCCAGACCACTCAGGAGTGGCCAGACCCA 1754
Qy 578 GAAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAACAGCCAAAGTAGCTTCAGC 637
Db 1753 GAAGAGGGCTCTTCAGCACCTAGAACCTGGAAGGGATGAACAGCCAAAGTAGCTTCAGC 1694
Qy 638 TTACAGCAGCCCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGACTGGATGGAGACA 697
Db 1693 TTACAGCAGCCCTCAAAATTTGGACCTGCCACCAGCTCCAGAGCTTGACTGGATGGAGACA 1634
Qy 698 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTTGGTAAACATCACTCA 757
Db 1633 GGACCATCTCTGACATTCATTGGCCATCAGGATATACAGGAGTTGGTAAACATCACTCA 1574
Qy 758 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCTGGGAAACCAAGAAATA 817
Db 1573 GGTGCCACACCTCCCTGGATGATCCAAAGATGAAGAATACATTGCTGGGAAACCAAGAAATA 1514
Qy 818 GGACCATCTCTGAGCAATTTCTTAAAGAAAGGAAACCAAGAGTTGAAAGAACTCCCC 877
Db 1513 GGACCATCTCTGAGCAATTTCTTAAAGAAAGGAAACCAAGAGTTGAAAGAACTCCCC 1454
Qy 878 CCAGACCGAGTTGGGGCCAACTTTGATCAGAGTCCAGGACCAAGTGGAGGCTGGCTGCC 937
Db 1453 CCAGACCGAGTTGGGGCCAACTTTGATCAGAGTCCAGGACCAAGTGGAGGCTGGCTGCC 1394
Qy 938 TCTTTTGGCGCGCTCTGGAATATGGAACGCGCTGGCAGTCCAGACATCAACTCCAAAA 997
Db 1393 TCTTTT-GGCGCGCTCTGGAATATGGAACGCGCTGGCAGTCCAGACATCAA-TTCAAAA 1336
Qy 998 CTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1057
Db 1335 CTGAAGCTGCAGCAATGAAGAGCAGTCAATACAGAAAAAGCTAATCATGCTCTCTAC 1276
Qy 1058 CAATACCATCAGCTTAAGCCAAAGTCAACCAACCCCTATTATACCTTCCACCCAAA 1117
Db 1275 CAATACCATCAGCTTAAGCCAAAGTCAACCAACCCCTATTATACCTTCCCA-CCAAA 1219
Qy 1118 TCTTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCAATTCATTGATTTAATAAGTT 1177
Db 1218 TCTTTTATCATGCTCTTTCTTAGGAAACAGACATCACTCAATTCATTGATTTAATAAGTT 1159
Qy 1178 TTATTTTC 1186
Db 1158 TTATTTTC 1150
```

Search completed: June 11, 2005, 04:01:07  
Job time : 3143.2 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:00 ; Search time 63.5671 Seconds  
(without alignments)  
2537.148 Million cell updates/sec

Title: US-09-155-676B-5  
Perfect score: 2190  
Sequence: 1 IGTVRRRSRIYDPRAXNS.....RGLESSLIPSTSGRGRTHF 417

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2178	99.5	417	2	AAW42401
2	795	36.3	1115	5	AAB71230
3	795	36.3	1115	7	ABW01535
4	795	36.3	1115	8	ADJ71905
5	670	30.6	156	4	ABG19580
6	554.5	25.3	741	4	ABG19581
7	503	23.0	312	4	ABG13680
8	309.5	14.1	65	4	AAM89643
9	198	9.0	104	2	AAV16779
10	196	8.9	104	7	ADA45163
11	111.5	5.1	850	4	ABB68702
12	111	5.1	838	4	ABB64668
13	110.5	5.0	1878	8	ADI28627
14	109	5.0	4952	5	ADH47759
15	109	5.0	4952	7	ADP68294
16	109	5.0	4952	8	ADL25642
17	109	5.0	5159	5	ADH48828
18	108.5	5.0	830	4	ABG08145
19	108.5	5.0	1225	3	AAW07973
20	108.5	5.0	1225	7	ADF09504
21	108.5	5.0	1233	4	ABG04996
22	107.5	4.9	1516	5	ABP69840
23	107	4.9	1419	5	ABP69842
24	107	4.9	1879	8	ADI28615
25	106	4.8	270	4	AAB68357

26	105.5	4.8	850	8	ABM80887	Abm80887 Tumour-as
27	105.5	4.8	852	8	ABM80886	Abm80886 Tumour-as
28	105.5	4.8	1942	7	ADE59894	Ades59894 Human pro
29	105.5	4.8	2246	4	ABG05850	Abg05850 Novel hum
30	105	4.8	366	8	ADR86145	Adr86145 Aspergill
31	105	4.8	669	4	ABB69682	Abb69682 Drosophil
32	105	4.8	1770	6	ABO14791	Abol14791 Novel hum
33	105	4.8	2701	6	ABR92087	AbR92087 Human cer
34	105	4.8	2819	4	AAB35408	Aab35408 Human 07C
35	105	4.8	2819	8	ADQ97653	Adq97653 Human can
36	104.5	4.8	526	6	ADA54293	Ada54293 Human pro
37	104.5	4.8	526	7	ADG31710	Adg31710 Human pro
38	104.5	4.8	1094	5	ABP73717	Abp73717 Candida a
39	104	4.7	475	8	ADR41765	Adr41765 Protein s
40	104	4.7	475	8	ADS14358	Adsl14358 Human 52-
41	103.5	4.7	578	2	AAW89273	Aaw89273 Granulocy
42	103.5	4.7	753	8	ADR14357	Adr14357 Human NF-
43	103.5	4.7	3111	4	ABB60327	Abb60327 Drosophil
44	103.5	4.7	3224	2	AAWS4235	Aaw54235 Human Nup
45	103.5	4.7	3224	6	ABO14790	Abol14790 Novel hum

ALIGNMENTS

RESULT 1  
AAW42401  
ID AAW42401 standard; protein; 417 AA.  
XX  
AC AAW42401;  
XX  
DT 15-APR-1998 (first entry)  
XX  
DE TRAF2 binding protein encoded by clone 15.  
XX  
KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;  
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;  
KW intracellular signalling activity; acute hepatitis;  
KW autoimmune-induced cell death.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 18 /note= "not specified"  
FT Misc-difference 320 /note= "not specified"  
FT Misc-difference 338 /note= "not specified"  
FT Misc-difference 356 /note= "not specified"  
FT Misc-difference 358 /note= "not specified"  
FT Misc-difference 388 /note= "not specified"  
FT Misc-difference 388 /note= "not specified"  
XX  
XX WO9737016-A1.  
XX  
XX 09-OCT-1997.  
XX  
XX 01-APR-1997; 97WO-IL000117.  
XX  
XX 02-APR-1996; 96IL-00117800.  
XX 26-AUG-1996; 96IL-00119133.  
XX (YEDA ) YEDA RES & DEV CO LTD.  
XX  
XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;  
XX WPI; 1997-503101/46.  
XX N-FSDB; AAV03325.  
XX  
XX DNA encoding tumour necrosis factor receptor-associated factor binding

PT molecule - used for modulation or mediation in cells of the activity of  
PT NF-KB.

XX Disclosure; Fig 5; 127pp; English.

XX The present sequence represents a TRAF2 binding protein, which is encoded  
CC by clone 15. Clone 15 is a partial clone, which lacks most of its 5' end  
CC of the coding DNA sequence. A cDNA library prepared from B-cells was  
CC screened for proteins that associate with TRAF2, and clone 15 was  
CC isolated. The clone 15 protein is capable of binding to at least amino  
CC acids 222-501 of TRAF2. The TRAF-2 binding proteins can be used for  
CC modulation or mediation in cells of the activity of NF-kappaB or any  
CC other intracellular signalling activity modulated or mediated by TRAF2.  
CC TRAF-binding proteins are especially used for prevention or treatment of  
CC pathological conditions associated with NF-kB induction, e.g. acute  
CC hepatitis, autoimmune-induced cell death, e.g. death of the beta  
CC cells of the pancreas that results in diabetes, the death of  
CC cells in graft rejection, the death of oligodendrocytes in the brain in  
CC multiple sclerosis, and AIDS-inhibited T cell suicide which causes  
CC proliferation of the AIDS virus and hence the AIDS disease. The proteins  
CC are also useful for screening of ligands capable of binding to a protein,  
CC which are useful for modulating cellular activity modulated/mediated by  
CC TRAF2

XX SQ Sequence 417 AA;

Query Match 99.5%; Score 2178; DB 2; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6.2e-190;  
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 IGVTWRSSRIVDPRAAXNSIRAHGPPFCGRGTTAASTORQLKEAFERLLPOVEAARK 60  
Db 1 IGVTWRSSRIVDPRAAXNSIRAHGPPFCGRGTTAASTORQLKEAFERLLPOVEAARK 60  
Qy 61 ATRAAQVERVYPHERCCWCLCGCEVREHLHGNTLVLYGGLLEHLASPEHKATNKF 120  
Db 61 ATRAAQVERVYPHERCCWCLCGCEVREHLHGNTLVLYGGLLEHLASPEHKATNKF 120  
Qy 121 WENKAEVOMKEFLVTPQDYARFKGMVKGDLSEYEEKDKVIKEMAAQIREVEQSRQV 180  
Db 121 WENKAEVOMKEFLVTPQDYARFKGMVKGDLSEYEEKDKVIKEMAAQIREVEQSRQV 180  
Qy 181 RSVLEPQAVDPPEGSSAPRSWKMNSQVASSLQQPSNLDLPAPELDMWETGPSITFIG 240  
Db 181 RSVLEPQAVDPPEGSSAPRSWKMNSQVASSLQQPSNLDLPAPELDMWETGPSITFIG 240  
Qy 241 HQDIPGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKOKLKLPPDRVGANF 300  
Db 241 HQDIPGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKOKLKLPPDRVGANF 300  
Qy 301 DHSRTSAGWLPSGPRLEXTWPLAVOTSTPKLQOXRSSHIQKKANHALYOLPXGXP 360  
Db 301 DHSRTSAGWLPSGPRLEXTWPLAVOTSTPKLQOXRSSHIQKKANHALYOLPXGXP 360  
Qy 361 KSTPLLYLPKFFIIFLRKQTSYFIFXFNKVLFPGLRGLESSLIPSTSGRGRTHF 417  
Db 361 KSTPLLYLPKFFIIFLRKQTSYFIFXFNKVLFPGLRGLESSLIPSTSGRGRTHF 417

RESULT 2

AAB71230

ID AAB71230 standard; protein; 1115 AA.

XX AAB71230;

AC AAB71230;

XX 18-NOV-2002 (first entry)

DE Human legless homologue hlgs-1 partial protein.

XX Legless; human; lgs; Wnt/Wingless signaling pathway; Wnt; Wg;  
KW tissue proliferation; tumour; cytoskeletal; cellular disorder; colon;  
KW blood disorder; cancer; breast; head and neck cancer; brain; thyroid;  
KW medulloblastoma; skin cancer; tissue regeneration; tissue repair.

XX OS

XX Homo sapiens.

XX US2002086986-A1.

XX PD 04-JUL-2002.

XX 27-JUL-2001; 2001US-00915543.

XX 28-JUL-2000; 2000US-0221502P.

XX (BASL/) BASLER K.

XX (BRUN/) BRUNNER E.

XX (PROE/) FROESCH B.

XX (KRAM/) KRAMPS T.

XX (PETE/) PETER O.

XX Basler K, Brunner E, Froesch B, Kramps T, Peter O;

XX WPI: 2002-635689/68.

XX N-PSDB; AAF88468.

Novel polypeptide useful in therapeutic method for treating disorders of cell fate such as cell differentiation or cell proliferation.

XX Example II; Fig 10B; 41pp; English.

XX This invention describes a novel polypeptide sharing one or more homologous amino acid domains with the legless (lgs) protein, a downstream component of the Wnt/Wingless (Wnt/Wg) signaling pathway involved in the formation and maintenance of spatial arrangements and proliferation of tissues during development, and in the formation and growth of many human tumours. The products of the invention have cytoskeletal activity and can be used to treat cellular disorders, blood disorders and cancers caused by over-stimulation of the Wnt pathway, where the cancerous condition is colon, breast, head and neck, brain, thyroid, medulloblastoma or skin cancer. The product could also be used to promote tissue regeneration and repair. This sequence represents the human legless (lgs) protein homologue hlgs-1 described in the disclosure of the invention

XX SQ Sequence 1115 AA;

Query Match 36.3%; Score 795; DB 5; Length 1115;

Best Local Similarity 55.6%; Pred. No. 7e-63;

Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

Qy 28 FFCGRGTTAASTORQLKEAFERLLPOVEAARKAIRAAQVERVYPHERCCWCLCGCEV 87

Db 28 FFCGRGTTAASTORQLKEAFERLLPOVEAARKAIRAAQVERVYPHERCCWCLCGCEV 87

Qy 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFVWENKAEVOMKEFLVTPQDYARFKSM 147

Db 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFVWENKAEVOMKEFLVTPQDYARFKSM 147

Qy 148 VGLGDSYEKEDKVIKEMAAQIREVEQSRQVRSVLEPQAVDPDEEGSSAPRSWKMNS 207

Db 148 VGLGDSYEKEDKVIKEMAAQIREVEQSRQVRSVLEPQAVDPDEEGSSAPRSWKMNS 207

Qy 948 VGLGDSYEKEDKVIKEMAAQIREVEQSRQVRSVLEPQAVDPDEEGSSAPRSWKMNS 990

Db 948 VGLGDSYEKEDKVIKEMAAQIREVEQSRQVRSVLEPQAVDPDEEGSSAPRSWKMNS 990

Qy 208 QVASSLQPSNLDLPAPELDMWETGPSITFIGHQDIPGVGNHSGATPPWMIQDEEYIA 267

Db 208 QVASSLQPSNLDLPAPELDMWETGPSITFIGHQDIPGVGNHSGATPPWMIQDEEYIA 267

Qy 991 -YALTVRSFAVLSR-----RTLKSGAFPPQ--TPEAHQ 1021

Db 991 -YALTVRSFAVLSR-----RTLKSGAFPPQ--TPEAHQ 1021

Qy 268 GNGEIGPSYEEFLKEKEKOKLKLPPDRVGANFPHSSRTSAGWLPSFGPRLEXTWPLAV 326

Db 268 GNGEIGPSYEEFLKEKEKOKLKLPPDRVGANFPHSSRTSAGWLPSFGPRLEXTWPLAV 326

Qy 1022 ARCLCAPRRGALKPEPPGFTLKGVPVPHHTTRKARPHAATSP-----RPRCTRQAP--- 1072

Db 1022 ARCLCAPRRGALKPEPPGFTLKGVPVPHHTTRKARPHAATSP-----RPRCTRQAP--- 1072

Qy 327 QHSTPKLKL--QOXRSSHIQKKA 347

Db 327 QHSTPKLKL--QOXRSSHIQKKA 347

Qy 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

Db 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 3



ABW01535  
ID ABW01535 standard; protein; 1115 AA.  
XX  
AC ABW01535;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Human lgs-1 protein.  
XX  
KW Legless protein; lgs; cell fate disorder; blood disease; gene therapy;  
KW cancer; tissue regeneration; tissue repair; cytostatic.  
XX  
OS Homo sapiens.  
XX  
FN US2003114413-A1.  
XX  
PD 19-JUN-2003.  
XX  
PF 19-DEC-2002; 2002US-00322579.  
XX  
PR 28-JUL-2000; 2000US-0221502P.  
PR 27-JUL-2001; 2001US-00915543.  
XX  
PA (UYZU-) UNIV ZURICH.  
XX  
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;  
XX WPI; 2003-829432/77.  
XX N-PSDB; AAD62643.  
XX  
PT Novel lgs polypeptide useful for isolation of lgs-binding proteins,  
PT diagnosing disorders of cell fate, treating diseases such as cancer.  
XX  
PS Claim 7; Fig 10B; Opp; English.  
XX  
XX The invention relates to novel legless (lgs) proteins and polynucleotides  
CC encoding such proteins. Lgs sequences are useful for the treatment of  
CC disorders of cell fate such as differentiation or proliferation. The  
CC invention is used to treat blood disease or a cancerous condition  
CC characterised by over-stimulation of the Wnt pathway such as colon,  
CC breast, head and neck, brain, thyroid, medulloblastoma or skin cancer and  
CC is administered to prevent progression from a pre-neoplastic or non-  
CC malignant condition to a neoplastic or malignant state. It is  
CC administered to promote tissue regeneration and repair. The invention is  
CC also useful in the therapy of diseases cost by an over-activation of Wg  
CC pathway. It is useful for reducing lgs gene expression in an invertebrate  
CC or vertebrate organism or an invertebrate or vertebrate cell line. The  
CC invention is also useful in gene therapy. The present sequence is human  
CC lgs-1 protein used in the invention  
XX  
SQ Sequence 1115 AA;  
Query Match 36.3%; Score 795; DB 7; Length 1115;  
Best Local Similarity 55.6%; Pred. No. 7e-63;  
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;  
QY 28 FFCGGTGTAASTQQLKAEFERLLPQVEAARKAIRAQVRYVPEHERCCWCLCCGCEV 87  
DB 828 FFCGGHYVSRKHQQLKEALLPQVEAARKAIRAQVRYVPEHERCCWCLCCGCEV 887  
QY 88 REHLHSHGNLTVLYGGLLEHLASPEHKKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 147  
DB 888 REHLHSHGNLTVLYGGLLEHLASPEHKKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 947  
QY 148 VKGLDSYEEDKDKYIKENAAQIREVQSRQEVRSVLEPQVPPPEGSSAPRSWKGNVS 207  
DB 948 VKGLDSYEEDKDKYIKENAAQIREVQSRQEVRSVLEPQVPPPEGSSAPRSWKGNVS 990  
QY 208 QVASSLOQPSNLDLPPAPELDWMETGSLTFIGHQDIPGVGNHSGATPPPMWIODEEYIA 267  
DB 991 -YALTVRSPAVLSR-----RFLKSGAFPPQ--TPEAHPQ 1021  
QY 268 GNOBIGPSYEEFLKEKEQKLLK-LPPDRVGANFDHSSRTSAGWLPSPGPRLEXWTPLAV 326

Db 1022 ARCLCAPRRGALKPEPPGRTLUKGVPHHTTRKARPHAATSP-----RPRCTROP--- 1072  
QY 327 QTSTPKLKL---QQXRSSHIOKKA 347  
Db 1073 -NKTQSLQLAGKARKKTALHLQTKA 1095  
RESULT 4  
ADJ71905  
ID ADJ71905 standard; protein; 1115 AA.  
XX  
AC ADJ71905;  
XX  
DT 20-MAY-2004 (first entry)  
XX  
DE Human Lgs/Bcl9 partial polypeptide.  
XX  
KW Human; legless; lgs; cell differentiation disorder;  
KW cell proliferation disorder; cancer; Wnt pathway; medulloblastoma; colon;  
KW breast; head; neck; brain; thyroid; skin; blood disease;  
KW tissue regeneration; tissue repair; cytostatic; Lgs/Bcl9.  
OS Homo sapiens.  
XX  
FN US2004038901-A1.  
XX  
PD 26-FEB-2004.  
XX  
PF 22-SEP-2003; 2003US-00664859.  
XX  
PR 28-JUL-2000; 2000US-0221502P.  
PR 27-JUL-2001; 2001US-00915543.  
XX  
PA (UYZU-) UNIV ZURICH.  
XX  
PI Basler K, Brunner E, Froesch B, Kramps T, Peter O;  
XX WPI; 2004-203288/19.  
XX N-PSDB; ADJ71904.  
XX  
PT Novel polypeptide sharing one or more homologue amino acid domains with  
PT legless protein being functional homologue of Legless, useful for  
XX diagnosing disorders of cell fate.  
XX Example 2; SEQ ID NO 17; 62pp; English.  
XX  
CC The invention relates to a polypeptide sharing one or more homologous  
CC amino acid domains with a Legless (Lgs) protein and is therefore a  
CC functional homologue of Lgs. The invention also relates to a nucleotide  
CC sequence encoding a protein present in invertebrate and/or vertebrate  
CC organisms, the nucleotide sequence encoding a protein comprising a  
CC positive function in a regulatory pathway and the use of the polypeptide  
CC for the isolation of lgs-binding proteins by carrying out an assay chosen  
CC from an in vitro binding assay with such a peptide or a co-  
CC immunoprecipitation from vertebrate or invertebrate cell lysates or a  
CC mammalian or yeast two hybrid assay. The polypeptide and polynucleotide  
CC are useful for treating disorders of cell fate, which involves  
CC administering therapeutic compounds chosen from invertebrate and  
CC vertebrate lgs protein homologues or fragments, antibodies, antibody  
CC fragments, lgs antisense DNA, lgs antisense RNA, lgs double-stranded RNA,  
CC small peptides or chemical and natural compounds being capable of  
CC interfering with Lgs function, synthesis and degradation. The disorders  
CC are related to cell differentiation or cell proliferation. The compound  
CC is administered to treat a cancerous condition by preventing progression  
CC from a pre-neoplastic or non-malignant condition to a neoplastic or  
CC malignant state. The cancerous condition is characterised by over-  
CC stimulation of the Wnt pathway and is medulloblastoma or cancer of the  
CC colon, breast, head and neck, brain, thyroid or skin. The therapeutic  
CC compound may also be administered to a blood disease to promote tissue  
CC regeneration and repair. This sequence represents a human Lgs/Bcl9  
XX partial polypeptide of the invention.

```
SQ Sequence 1115 AA;
Query Match 36.3%; Score 795; DB 8; Length 1115;
Best Local Similarity 55.6%; Pred. No. 7e-63;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGTTTAASTORQLKEAFERLLPQVEAAKRAIRAAQVRYVPEHERCCWCLCCGCEV 87
DB 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAAKRAIRAAQVRYVPEHERCCWCLCCGCEV 887
QY 88 REHLHGNTLVLYGGLLEHLASPEHKATNKFWENKAEVOMKEKFLVTPDYARFKKSM 147
DB 888 REHLHGNTLVLYGGLLEHLASPEHKATNKFWENKAEVOMKEKFLVTPDYARFKKSM 947
QY 148 VGLGDSYEBEKDKVKEMAAQIREVEQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
DB 948 VGLGDSYEBEKDKVKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990
QY 208 QVASSLQSPNSLDLPAPPELDWMTGSLTFIGHQDIPGVGNIHSGATPPWMIQDEEYIA 267
DB 991 -VALTVRSPAVLSR-----RTLKSAGAPPQ--TPEAHPQ 1021
QY 268 GNGEIQSPSYEEFLKEKEKQK-LKLPDRVGANFDHSSRTSAGWLPSPGPRLEXWTPPLAV 326
DB 1022 ARCLCAPRRGALKPEPPGRTLKLGVPDHTTRKARPHAAKTSP-----RPRCTROAP--- 1072
QY 327 QTSTPKLKL---QOXRSSHIOKKA 347
DB 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 5
ABGI19580
ID ABGI19580 standard; protein; 156 AA.
XX AC ABGI19580;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19571.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS83767.
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 49939; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
```

```
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 156 AA;
Query Match 30.6%; Score 670; DB 4; Length 156;
Best Local Similarity 77.9%; Pred. No. 1e-52;
Matches 134; Conservative 4; Mismatches 12; Indels 22; Gaps 4;

QY 6 WRRSRIVDPRAAXNSIRAHEGPFPCGRGTFTAASTORQLKEAFERLLPQVEAAKRAIRAA 65
DB 5 WSRARL-QPQAPA-----AARGGF--GEAPAPGA-----EVEAAKRAIRAA 42
QY 66 QVERYVPEHERCCWCLCCGCEVREHLHGNTLVLYGGLLEHLASPEHKATNKFWENKA 125
DB 43 QVERYVPEHERCCWCLCCGCEVREHLHGNTLVLYGGLLEHLASPEHKATNKFWENKA 102
QY 126 EVOMKEKFLVTPDYARFKKSMVKGGLDSYEBEKDKVKEMAAQIREVEQSRQ 177
DB 103 EVOMKEKFLVTPDYARFKKSMVKGGLDSYEBEKDKVKEMAAQIREVEQSRQ 154

RESULT 6
ABGI19581
ID ABGI19581 standard; protein; 741 AA.
XX AC ABGI19581;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #19572.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS83768.
XX XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
```

```
PS Claim 20; SEQ ID NO 49940; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 741 AA;
SQ
Query Match 25.3%; Score 554.5; DB 4; Length 741;
Best Local Similarity 40.7%; Pred. No. 3.7e-41;
Matches 135; Conservative 18; Mismatches 52; Indels 127; Gaps 11;
Qy 66 QERYVPEHERCCWCLCGCEVREHLSHGNTLVYGLLEHL-----ASPEHKK----- 114
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 175
175 KVERYVPEHERCCWCLCGCEVREHLSHGNTLVYGLLEHLARTASIVAAPNKKQFRCFL 234
Qy 115 ATNKFWEKAEVQMKFLVTPDYARFKSKVMKGLSDSYEEKDKVKEKMAAQIREVBQ 174
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 235
235 STGK--WINNG-----IVFQEDRI-----IDYDTDES-----QK 263
Qy 175 SRQEVRSVLBPQAVPDPEGSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMWETGP 234
Db |||||-----SR----- 265
Qy 235 SLTFIGHQDIPGVGNIHSGATPPWMIQDEYIAGNQEI GPSYEFLEKEKEKQKLKLPDP 294
Db |||||-----PG-----RRGATPPWMIQDEYIAGNQEI GPSYEFLEKEKEKQKLKLPDP 311
Qy 295 RVGANFDHSSRTSAGWLPFGPRLEXWTPLAVQTSTPKLKLOQRSSHIQKANHLYQL 354
Db |||||-----RVGN-----RVWN-----NGRRWQSRAPEGEREVPWSEBQT 359
Qy 355 PXGKKPKSTK-PLLYLPPKFFIIVFLRKQTVS 385
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||: 360
360 PDPTSLRLRTPAEVAPRSKRCVLERKQPIYS 391
RESULT 7
ABG13680
ID ABG13680 standard; protein; 312 AA.
XX
XX ABG13680;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #13671.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
FN
XX 11-OCT-2001.
PD
```

---

```
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AAS77867.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 44039; 103pp; English.
FS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 312 AA;
SQ
Query Match 23.0%; Score 503; DB 4; Length 312;
Best Local Similarity 49.3%; Pred. No. 5.2e-37;
Matches 113; Conservative 14; Mismatches 38; Indels 64; Gaps 6;
Qy 40 TQRLKEAFERL-----LPQVEAARAKAIRAAQVERYPVEHERCCWCLCGCEVREHL 91
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 108
108 TQRLPAADWMPAATAITGITDVEAARAKAIRAAQVERYPVEHERCCWCLCGCEVREHL 167
Qy 92 SHGNLTVLVYGLLEHLASPEHKATNKFWMENKAEVQMKFLVTPDYARF----- 143
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 168
168 SHGNLTVLVYGLLEHLASPEHKATNKFWMENKAEVQMKFLVTPDYARSHLYGTGPMG 227
Qy 144 ----KKSVMVGL-----DSYREKEDKVIKMAAQIREVQSROEVRSV-LEPQA 188
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||: 228
228 SAVPSQACLEALLIIVPWGACGVSQEEES-----PAGSKDEPGEQVELKEEA 276
Qy 189 VPDPEGSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMWETGPSIT 237
Db :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||: 277
277 EAPVEDGSGPP-----PPEPKGDATPGEKAT 303
RESULT 8
AAM89643
ID AAM89643 standard; protein; 65 AA.
XX
XX AAM89643;
AC
XX
XX 07-NOV-2001 (first entry)
DT
XX
```

DE	Human immune/haematopoietic antigen SEQ ID NO:17236.	PR	21-SEP-2000; 2000US-0234223P.
XX		PR	21-SEP-2000; 2000US-0234274P.
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;	PR	25-SEP-2000; 2000US-0234997P.
KW	cystostatic; gene therapy; vaccine; metastasis.	PR	25-SEP-2000; 2000US-0234998P.
XX		PR	26-SEP-2000; 2000US-0235484P.
OS	Homo sapiens.	PR	27-SEP-2000; 2000US-0235834P.
XX		PR	27-SEP-2000; 2000US-0235836P.
XX	WO200157182-A2.	PR	29-SEP-2000; 2000US-0236327P.
XX		PR	29-SEP-2000; 2000US-0236367P.
XX		PR	29-SEP-2000; 2000US-0236368P.
PD		PR	29-SEP-2000; 2000US-0236369P.
XX		PR	29-SEP-2000; 2000US-0236370P.
PF	17-JAN-2001; 2001WO-US001354.	PR	02-OCT-2000; 2000US-0236802P.
XX		PR	02-OCT-2000; 2000US-0237037P.
PR	31-JAN-2000; 2000US-0179065P.	PR	02-OCT-2000; 2000US-0237038P.
PR	04-FEB-2000; 2000US-0180628P.	PR	02-OCT-2000; 2000US-0237039P.
PR	24-FEB-2000; 2000US-0184664P.	PR	13-OCT-2000; 2000US-0239935P.
PR	02-MAR-2000; 2000US-0186350P.	PR	13-OCT-2000; 2000US-0239937P.
PR	16-MAR-2000; 2000US-0189874P.	PR	20-OCT-2000; 2000US-0240960P.
PR	17-MAR-2000; 2000US-0190076P.	PR	20-OCT-2000; 2000US-0241221P.
PR	18-APR-2000; 2000US-0198123P.	PR	20-OCT-2000; 2000US-0241785P.
PR	19-MAY-2000; 2000US-0205515P.	PR	20-OCT-2000; 2000US-0241786P.
PR	07-JUN-2000; 2000US-0209467P.	PR	20-OCT-2000; 2000US-0241787P.
PR	28-JUN-2000; 2000US-0214886P.	PR	20-OCT-2000; 2000US-0241808P.
PR	30-JUN-2000; 2000US-0215135P.	PR	20-OCT-2000; 2000US-0241809P.
PR	07-JUL-2000; 2000US-0216647P.	PR	20-OCT-2000; 2000US-0241826P.
PR	07-JUL-2000; 2000US-0216880P.	PR	01-NOV-2000; 2000US-0244617P.
PR	11-JUL-2000; 2000US-0217487P.	PR	08-NOV-2000; 2000US-0246474P.
PR	11-JUL-2000; 2000US-0217496P.	PR	08-NOV-2000; 2000US-0246475P.
PR	14-JUL-2000; 2000US-0218290P.	PR	08-NOV-2000; 2000US-0246476P.
PR	26-JUL-2000; 2000US-0220963P.	PR	08-NOV-2000; 2000US-0246477P.
PR	26-JUL-2000; 2000US-0220964P.	PR	08-NOV-2000; 2000US-0246478P.
PR	14-AUG-2000; 2000US-0224518P.	PR	08-NOV-2000; 2000US-0246523P.
PR	14-AUG-2000; 2000US-0224519P.	PR	08-NOV-2000; 2000US-0246524P.
PR	14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000; 2000US-0246525P.
PR	14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000; 2000US-0246526P.
PR	14-AUG-2000; 2000US-0225266P.	PR	08-NOV-2000; 2000US-0246527P.
PR	14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000; 2000US-0246528P.
PR	14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000; 2000US-0246532P.
PR	14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000; 2000US-0246609P.
PR	14-AUG-2000; 2000US-0225447P.	PR	08-NOV-2000; 2000US-0246610P.
PR	14-AUG-2000; 2000US-0225757P.	PR	08-NOV-2000; 2000US-0246611P.
PR	14-AUG-2000; 2000US-0225758P.	PR	08-NOV-2000; 2000US-0246613P.
PR	14-AUG-2000; 2000US-0225759P.	PR	17-NOV-2000; 2000US-0249207P.
PR	18-AUG-2000; 2000US-0226279P.	PR	17-NOV-2000; 2000US-0249208P.
PR	22-AUG-2000; 2000US-0226681P.	PR	17-NOV-2000; 2000US-0249209P.
PR	22-AUG-2000; 2000US-0226686P.	PR	17-NOV-2000; 2000US-0249210P.
PR	22-AUG-2000; 2000US-0227182P.	PR	17-NOV-2000; 2000US-0249211P.
PR	23-AUG-2000; 2000US-0227009P.	PR	17-NOV-2000; 2000US-0249212P.
PR	30-AUG-2000; 2000US-0228924P.	PR	17-NOV-2000; 2000US-0249213P.
PR	01-SEP-2000; 2000US-0229287P.	PR	17-NOV-2000; 2000US-0249214P.
PR	01-SEP-2000; 2000US-0229343P.	PR	17-NOV-2000; 2000US-0249215P.
PR	01-SEP-2000; 2000US-0229344P.	PR	17-NOV-2000; 2000US-0249216P.
PR	01-SEP-2000; 2000US-0229345P.	PR	17-NOV-2000; 2000US-0249217P.
PR	05-SEP-2000; 2000US-0229509P.	PR	17-NOV-2000; 2000US-0249218P.
PR	05-SEP-2000; 2000US-0229513P.	PR	17-NOV-2000; 2000US-0249244P.
PR	06-SEP-2000; 2000US-0230437P.	PR	17-NOV-2000; 2000US-0249245P.
PR	06-SEP-2000; 2000US-0230438P.	PR	17-NOV-2000; 2000US-0249264P.
PR	08-SEP-2000; 2000US-0231242P.	PR	17-NOV-2000; 2000US-0249265P.
PR	08-SEP-2000; 2000US-0231243P.	PR	17-NOV-2000; 2000US-0249297P.
PR	08-SEP-2000; 2000US-0231244P.	PR	17-NOV-2000; 2000US-0249299P.
PR	08-SEP-2000; 2000US-0231413P.	PR	17-NOV-2000; 2000US-0249300P.
PR	08-SEP-2000; 2000US-0231414P.	PR	01-DEC-2000; 2000US-0250160P.
PR	08-SEP-2000; 2000US-0232080P.	PR	01-DEC-2000; 2000US-0250391P.
PR	08-SEP-2000; 2000US-0232081P.	PR	05-DEC-2000; 2000US-0251030P.
PR	12-SEP-2000; 2000US-0231968P.	PR	05-DEC-2000; 2000US-0251988P.
PR	14-SEP-2000; 2000US-0232397P.	PR	05-DEC-2000; 2000US-0256719P.
PR	14-SEP-2000; 2000US-0232398P.	PR	06-DEC-2000; 2000US-0251479P.
PR	14-SEP-2000; 2000US-0232399P.	PR	08-DEC-2000; 2000US-0251856P.
PR	14-SEP-2000; 2000US-0232400P.	PR	08-DEC-2000; 2000US-0251868P.
PR	14-SEP-2000; 2000US-0232401P.	PR	08-DEC-2000; 2000US-0251869P.
PR	14-SEP-2000; 2000US-0233063P.	PR	08-DEC-2000; 2000US-0251989P.
PR	14-SEP-2000; 2000US-0233064P.		
PR	14-SEP-2000; 2000US-0233065P.		

```

PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX N-PSDB; AAK62424.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Claim 11; SEQ ID NO 17236; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention
XX
XX SQ Sequence 65 AA;

Query Match 14.1%; Score 309.5; DB 4; Length 65;
Best Local Similarity 92.2%; Pred. No. 2.5e-20;
Matches 59; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

Qy 238 FIGH---QDIPGVGNHSGATPPWMIQDEEVIAGNQGIGPSYEFLEKEKQKLLKLPD 294
Db 2 FFHFQQLQDIPGVGNHSGATPPWMIQDEEVIAGNQGIGPSYEFLEKEKQKLLKLPD 61

Qy 295 RVGA 298
Db 62 RVGA 65

RESULT 9
AAI16779
ID AAI16779 standard; protein; 104 AA.
XX
XX AAI16779;
XX
XX 27-JUL-1999 (first entry)
XX
XX Human secreted protein (clone as20_2).
XX
XX Secreted protein; human; tissue marker; genetic disease; gene therapy;
XX veterinary medicine; cell proliferation; immunostimulant; infection;
XX immunosuppressant; autoimmune disease; organ rejection; tumour; anaemia;
XX haematopoiesis; wound healing; fertility control; chemotaxis; analgesic;
XX thrombolytic; haemophilia; infarction; antimicrobial agent; cancer.
XX
XX Homo sapiens.
XX
XX WO9924469-A1.
XX
XX 20-MAY-1999.
XX
XX 06-NOV-1998; 98WO-US023829.
XX

```

---

```

PR 07-NOV-1997; 97US-00965789.
PR 04-NOV-1998; 98US-00185936.
XX
XX (GEMY ) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Evans C, Merberg D;
XX Treacy M, Agostino MJ;
XX
XX WPI; 1999-327362/27.
XX N-PSDB; AAX60579.
XX
XX Nucleic acid encoding secreted human proteins.
XX
XX Claim 8; Page 92; 107pp; English.
XX
XX The invention provides polynucleotides (AAX60579-X60687) encoding
XX specific secreted human proteins (AAY16779-Y16787). The nucleic acid
XX sequences are deposited under the accession number ATCC 98580. The
XX polynucleotides are used as tissue markers, chromosomal tags, for
XX diagnosis of genetic diseases, to generate anti-protein or anti-DNA
XX antibodies, also as nutritional sources and supplements and in gene
XX therapy. The secreted proteins are useful therapeutically, in human or
XX veterinary medicine, e.g. for modulating cell proliferation or
XX differentiation, as immunostimulants or immunosuppressants (for treating
XX infections, autoimmune disease, organ rejection, or to induce tumour
XX immunity), as regulators of haematopoiesis (e.g. for treating anaemia or
XX in conjunction with tumour therapy), to stimulate growth of tissue for
XX wound healing, as fertility control agents, for regulating chemotaxis or
XX chemokines (e.g. for directing cells to tumours or sites of infection), as
XX haemostatic and thrombolytic agents (e.g. in treatment of haemophilia or
XX infarctions), as antimicrobial agents, for modifying biorhythms,
XX appetite, or metabolism, as analgesics and many other uses. The proteins
XX are also used to raise antibodies, used as diagnostic immunoassay
XX reagents also (when neutralizing) for treating e.g. cancer
XX
XX SQ Sequence 104 AA;

Query Match 9.0%; Score 198; DB 2; Length 104;
Best Local Similarity 62.5%; Pred. No. 7.9e-10;
Matches 40; Conservative 5; Mismatches 15; Indels 4; Gaps 1;

Qy 201 SWKGMNSQVASSLQPPNLDLPPAPELDWMETGSLTFIGHQDIPGVGNHSGATPPWMI 260
Db 19 AWMFFSQVASSLQPPNLDLPPAPELDWMETGSLTFIGHQ----VQRISKPEEGQWSL 74

Qy 261 QDEE 264
Db 75 RSQD 78

RESULT 10
ADA45163
ID ADA45163 standard; protein; 104 AA.
XX
XX ADA45163;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human polypeptide #69.
XX
XX Human; genetic disorder; genetic fingerprinting; autoimmune disorder;
XX multiple sclerosis; systemic lupus erythematosus; graft-versus-host disease; anaemia;
XX insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;
XX periodontal disease; bone fracture; cartilage damage;
XX central nervous system disorder; Alzheimer's disease;
XX Parkinson's disease; cancer; nutrition; carbon source; nitrogen source;
XX carbohydrate source.
XX
XX Homo sapiens.
XX
XX US2003044935-A1.
XX
XX 06-MAR-2003.
XX

```

```

XX PF 21-DEC-2000; 2000US-00746783.
XX PF 11-JUN-1997; 97US-0086236P.
XX PR 12-JUN-1997; 97US-0086234P.
XX PR 08-JUL-1997; 97US-0092115P.
XX PR 08-SEP-1997; 97US-0093045P.
XX PR 02-OCT-1997; 97US-0090100P.
XX PR 27-OCT-1997; 97US-00958304.
XX PR 07-NOV-1997; 97US-0090111P.
XX PR 05-JUN-1998; 98US-00092722.
XX PR 11-JUN-1998; 98US-00096287.
XX PR 17-JUN-1998; 98US-00098588.
XX PR 04-AUG-1998; 98US-00130189.
XX PR 08-SEP-1998; 98US-00149633.
XX PR 01-OCT-1998; 98US-00165960.
XX PR 04-NOV-1998; 98US-00185936.
XX (JACO/) JACOBS K.
XX PA (MCCO/) MCCOY J M.
XX PA (LVAL/) LA VALLIE E R.
XX PA (COLL/) COLLINS-RACIE L A.
XX PA (EVAN/) EVANS C.
XX PA (MERB/) MERBERG D.
XX PA (TREA/) TREACY M.
XX PA (SPAU/) SPAULDING V.
XX
XX PI Jacobs K, McCoy JM, La Vallie ER, Collins-Racie LA, Evans C;
XX PI Merberg D, Treacy M, Spaulding V;
XX DR N-PSDB; ADA45162.
XX
XX PT New polypeptides and polynucleotides having biological activities, useful
XX PT as nutritional sources or supplements, or for treating e.g. autoimmune
XX PT diseases, cancers, bone fractures or damages, or central nervous system
XX PT disorders.
XX PS Disclosure; Page 238-239; 288pp; English.
XX
XX CC The invention relates to human polynucleotides and the polypeptides they
XX CC encode. The polynucleotides can be used to express recombinant proteins
XX CC for analysis, characterisation or therapeutic use, as markers for tissues
XX CC in which the corresponding protein is expressed, as molecular weight
XX CC markers on Southern gels, as chromosome markers or tags to identify
XX CC chromosomes or to map related gene positions, to compare with endogenous
XX CC DNA sequences in patients to identify potential genetic disorders, as
XX CC probes to hybridise and discover novel related DNA sequences, as a source
XX CC of information to derive PCR primers for genetic fingerprinting, to raise
XX CC anti-protein antibodies and in gene therapy. The proteins can be used to
XX CC raise antibodies or to elicit another immune response, as reagents in
XX CC assays designed to quantitatively determine levels of the protein in
XX CC biological fluids, as markers for tissues in which the corresponding
XX CC protein is preferentially expressed and to treat autoimmune disorders
XX CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
XX CC diabetes mellitus or graft-versus-host disease), anaemias, periodontal
XX CC diseases, bone fractures, cartilage damage, central nervous system
XX CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
XX CC The proteins and polynucleotides are also useful as nutritional sources
XX CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
XX CC sequence represents a human polypeptide of the invention.
XX SQ Sequence 104 AA;
XX
XX Query Match 8.9%; Score 196; DB 7; Length 104;
XX Best Local Similarity 60.9%; Pred. No. 1.2e-09;
XX Matches 39; Conservative 6; Mismatches 15; Indels 4; Gaps 1;
XX
XX QY 201 SWKGMNSQVASSLQPSNLDLPPAPELDWMETGPSLTFTIGHQDIPGVGNHSGATPPWMI 260
XX :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 19 AWMFFSQVASSLQPSNLDLPPAPELDWMETGPSLTFTIGHQ----VQRISKPEGQWSL 74
XX
XX QY 261 QDEE 264
XX
Db 21-DEC-2000; 2000US-00746783.
Db 11-JUN-1997; 97US-0086236P.
Db 12-JUN-1997; 97US-0086234P.
Db 08-JUL-1997; 97US-0092115P.
Db 08-SEP-1997; 97US-0093045P.
Db 02-OCT-1997; 97US-0090100P.
Db 27-OCT-1997; 97US-00958304.
Db 07-NOV-1997; 97US-0090111P.
Db 05-JUN-1998; 98US-00092722.
Db 11-JUN-1998; 98US-00096287.
Db 17-JUN-1998; 98US-00098588.
Db 04-AUG-1998; 98US-00130189.
Db 08-SEP-1998; 98US-00149633.
Db 01-OCT-1998; 98US-00165960.
Db 04-NOV-1998; 98US-00185936.
Db (JACO/) JACOBS K.
Db PA (MCCO/) MCCOY J M.
Db PA (LVAL/) LA VALLIE E R.
Db PA (COLL/) COLLINS-RACIE L A.
Db PA (EVAN/) EVANS C.
Db PA (MERB/) MERBERG D.
Db PA (TREA/) TREACY M.
Db PA (SPAU/) SPAULDING V.
Db
Db PI Jacobs K, McCoy JM, La Vallie ER, Collins-Racie LA, Evans C;
Db PI Merberg D, Treacy M, Spaulding V;
Db DR N-PSDB; ADA45162.
Db
Db PT New polypeptides and polynucleotides having biological activities, useful
Db PT as nutritional sources or supplements, or for treating e.g. autoimmune
Db PT diseases, cancers, bone fractures or damages, or central nervous system
Db PT disorders.
Db PS Disclosure; Page 238-239; 288pp; English.
Db
Db CC The invention relates to human polynucleotides and the polypeptides they
Db CC encode. The polynucleotides can be used to express recombinant proteins
Db CC for analysis, characterisation or therapeutic use, as markers for tissues
Db CC in which the corresponding protein is expressed, as molecular weight
Db CC markers on Southern gels, as chromosome markers or tags to identify
Db CC chromosomes or to map related gene positions, to compare with endogenous
Db CC DNA sequences in patients to identify potential genetic disorders, as
Db CC probes to hybridise and discover novel related DNA sequences, as a source
Db CC of information to derive PCR primers for genetic fingerprinting, to raise
Db CC anti-protein antibodies and in gene therapy. The proteins can be used to
Db CC raise antibodies or to elicit another immune response, as reagents in
Db CC assays designed to quantitatively determine levels of the protein in
Db CC biological fluids, as markers for tissues in which the corresponding
Db CC protein is preferentially expressed and to treat autoimmune disorders
Db CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
Db CC diabetes mellitus or graft-versus-host disease), anaemias, periodontal
Db CC diseases, bone fractures, cartilage damage, central nervous system
Db CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
Db CC The proteins and polynucleotides are also useful as nutritional sources
Db CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
Db CC sequence represents a human polypeptide of the invention.
Db SQ Sequence 104 AA;
Db
Db Query Match 8.9%; Score 196; DB 7; Length 104;
Db Best Local Similarity 60.9%; Pred. No. 1.2e-09;
Db Matches 39; Conservative 6; Mismatches 15; Indels 4; Gaps 1;
Db
Db QY 201 SWKGMNSQVASSLQPSNLDLPPAPELDWMETGPSLTFTIGHQDIPGVGNHSGATPPWMI 260
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 19 AWMFFSQVASSLQPSNLDLPPAPELDWMETGPSLTFTIGHQ----VQRISKPEGQWSL 74
Db
Db QY 261 QDEE 264
Db

```

QY 275 SYEFLKEKEKQKLLPPDRVGANFDSHRSRTSAGWLPSF-GPRLEXWTPILAVQTS----- 329  
Db 459 EDDPNTNYAEQVIKKTTLVEEDDDFDPRABEHATEPPFLAAPQRDILLAGSATDLSQVVP 518  
QY 330 ---TPKLLQXRSRH--IQKANHALYQLPGKXKPKSTKPLLYLPKPKFIIVFLRKQTY 384  
Db 519 APLAPTLSDQEAEDFDPDTSASVALVQ-----PKSTE-----LRFLERELL 561  
QY 385 SFIXFNKVLFFGLRG--LESSLS 405  
Db 562 N-----YSLDGVTLKHSLS 576  
RESULT 12  
ID ABB64668  
XX ABB64668 standard; protein; 838 AA.  
AC ABB64668;  
XX 26-MAR-2002 (first entry)  
DT 26-MAR-2002 (first entry)  
DE Drosophila melanogaster polypeptide SEQ ID NO 20796.  
XX Drosophila: developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX Drosophila melanogaster.  
OS WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
PA Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
DR N-PSDB; ABL08771.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX Disclosure; SEQ ID NO 20796; 21pp + Sequence Listing; English.  
PS The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 838 AA;  
Query Match 5.1%; Score 111; DB 4; Length 838;  
Best Local Similarity 23.7%; Pred. No. 1.5;  
Matches 33; Conservative 21; Mismatches 39; Indels 18; Gaps 4;  
QY 152 DSYEEKEDKVIKEMAAQIREVQSRQEVRSVLE---PQAVDPPEGSSAPRSWKGMNSQ 208  
Db 239 ESKDAKEIIVQVEKTERDPPDKAKVEKTIKVKIPVDSPEAEENSPPSGIPPPSQ 298  
QY 209 VASSLQOOPSNL--DLPPAPELDMWMTGSLTFIGHQDIPGVGNIHSGATPP 257

Db 299 LPSDIPPPSQSPDIPPSQL-----PS-----DIPPPSQSPSGTPPP 336  
RESULT 13  
ID ADI28627  
XX ADI28627 standard; protein; 1878 AA.  
AC ADI28627;  
XX 22-APR-2004 (first entry)  
DT 22-APR-2004 (first entry)  
DE Mycoplasma hyopneumoniae immunogen protein C28-MHP545.  
XX Pneumonia; vaccine; diagnosis; C28-MHP545; immunogen; immunostimulant;  
KW antiinflammatory.  
XX Mycoplasma hyopneumoniae.  
OS WO2004003161-A2.  
XX 08-JAN-2004.  
PD 27-JUN-2003; 2003WO-US020460.  
XX 28-JUN-2002; 2002US-0392632P.  
PR (IOWA ) UNIV IOWA STATE RES FOUND INC.  
PA (NSWA-) NSW AGRIC.  
XX Minion CF, Mahairas GG, Djordjevic SP;  
XX WPI; 2004-083044/08.  
DR N-PSDB; ADI28626.  
XX New immunogenic Mycoplasma hyopneumoniae polypeptide, useful in eliciting  
PT an immune response and in treating or preventing enzootic pneumonia.  
XX Claim 11; SEQ ID NO 20; 81pp; English.  
XX The present sequence is the protein sequence of C28-MPH545, an  
CC immunogenic P102 paralog from Mycoplasma hyopneumoniae strain J. The  
CC invention provides M. hyopneumoniae polypeptides and nucleic acids  
CC ADI28608-ADI28627, including C28-MPH545, as well as vectors and host  
CC cells. Compositions containing M. hyopneumoniae polypeptides and nucleic  
CC acids are provided for use in methods of treating pigs to prevent  
CC enzootic pneumonia. Diagnostic tests for detecting M. hyopneumoniae  
CC infection in swine herds are also provided.  
XX  
SQ Sequence 1878 AA;  
Query Match 5.0%; Score 110.5; DB 8; Length 1878;  
Best Local Similarity 23.7%; Pred. No. 5.6;  
Matches 71; Conservative 45; Mismatches 132; Indels 51; Gaps 14;  
QY 92 SHGNLTLYLGGLEHL-----ASPEHKKATNKFWMENKAEVQMKFKL--VTPQDYA 141  
Db 414 SLGYNFLFDLASHLDYTLFVSKAKIKQSSITKKLFELPIKISLSKSSILGDQEPNIKT 473  
QY 142 RPKKSMVKGLDSYEEKB-DKVIKEMAAQI-REVEQSRQEVRSVLEPQAVDPPEGS-SA 198  
Db 474 LFEKEVTFKLDNFRDVEIEKAFGLLYPGVNEELQARREQRASLEKEKAKGLKEFSSQK 533  
QY 199 PRSWKGMNSQVASSLQOOPSNL--LPPAPELDMWMTGSLTFIGHQDIPGVGNIHSGATP 256  
Db 534 DENUKAINNQ--DGLSEDDNITERLPENSIQY-----QBERAGLUG---SSDPK 577  
QY 257 PMWIQDEE---YIAGNQEIGPSEYEEFLKEKEKQKLLPPDRVGANFDSHRSRTSAGWLP 312  
Db 578 PYMKIVQONQRYVLAKSQ-----IQELIKAKDYTKLAKLLSNRHTYINISLLEKEQ---LF 629  
QY 313 SFGRLXEWTFPLA-----VQTSPTKLQOQXRS-----HIQKANHALYQLPGXKKPKST 363

CC disease, Parkinson's disease, epilepsy, immune disorders  
CC (osteoarthritis), haematopoietic disorders, inflammatory skin disorders,  
CC asthma, and various dyslipidaemias. NOV18 comprises a ARL-like protein  
CC and maps to chromosome 12q12-q14.  
XX  
SQ Sequence 4952 AA;  
  
Query Match 5.0%; Score 109; DB 5; Length 4952;  
Best Local Similarity 20.9%; Pred. No. 32;  
Matches 80; Conservative 51; Mismatches 149; Indels 102; Gaps 15;  
  
Qy 101 GGLLEHLASPE-----HKATNKFWNKA-----EVQMKKFLVTPQDYAR 142  
Db 2609 GGPAPHLTSPSLSPGSGSLLEKFELESGALTPGCPAASGDELDMESSLVASE---- 2664  
  
Qy 143 FKSMVKGLDSYEKKDKVTKEMAAQIREVEQSRQEVRSVLPE-----QAVDPDEGSS 197  
Db 2665 -LPFLLEHLEHKKELQKKQLSAQLQPAQQQQQQQQHSLLPAPGPAQAMSLPHEGSS 2723  
  
Qy 198 APRSWKGMNSQVA-----SSLQOFSNLDLPPAPELDMMETGSPSLTFIGHQ----- 242  
Db 2724 P--SLAGSQQLSLGLAVARQPLQPLMPTQPPAHALQ--QLAPSNMAMVSNQGHMLSGQ 2780  
  
Qy 243 -----DIPGVGNHSGATPPWMIQDEYIAGNQEIGPSY-----EEF 279  
Db 2781 HGGQAGLVPOQSSQPVLSQKPMGTMPFSCMKPQQLAMQOQLANSFPDDTDLKFAEDI 2840  
  
Qy 280 LKEKEKOK-----LKKLPDRVGCANPDHSSRTSAGWLPSPGPRLEXWTFPLAVQTSTPKL 333  
Db 2841 IGPIAKAKMVALKGIKV-----MAQSGIVAPGWNQ-----QVSLLAQ 2880  
  
Qy 334 KLOQXRSRSHIQKCANHALYQLPXGKPKSTKPLLYLPKPF--IIVFLRKQTSFXIFXN- 390  
Db 2881 RLSCGPSSDLQNHVAAGSGQERSAGDPSQPRP---NPPTFAQGVINEADQORQEWLWLFHT 2937  
  
Qy 391 -KVLFPFGLRGLESSLIPSTSR 411  
Db 2938 QQLQOMQLKVLLEQIGVHRKSR 2959  
  
RESULT 15  
ADP68294  
ID ADP68294 standard; protein; 4952 AA.  
XX  
AC ADP68294;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human NOV18a protein, an ARL like protein SeqID 56.  
XX  
KW human; NOX; Alzheimer's disease; Huntington's; inflammatory;  
KW Crohn's disease; rheumatoid arthritis; immunological; endocrine;  
KW pigmentation; haematopoietic; psychotic; autoimmune; muscular;  
KW osteoporosis; angina pectoris; hypotension; anxiety; bulimia;  
KW cancer; manic depression; viricide; antibacterial; analgesic;  
KW neuroprotective; neurotropic; cerebroprotective; anticonvulsant;  
KW dermatological; osteopathic; antiarthritic; antiinflammatory; cytostatic;  
KW hypotensive; cardiac; hypertensive; antiulcer; antiallergic;  
KW antianginal; immunosuppressive; antidepressant; neurodegenerative; ARL.  
XX Homo sapiens.  
XX WO200281510-A2.  
XX  
XX 17-OCT-2002.  
XX  
XX 18-JAN-2002; 2002WO-US001467.  
XX  
XX 18-JAN-2001; 2001US-0262454P.  
PR 23-JAN-2001; 2001US-0263605P.  
PR 25-JAN-2001; 2001US-0264159P.  
PR 31-JAN-2001; 2001US-0265517P.  
PR 07-FEB-2001; 2001US-0267057P.

Db 630 EVNPRIPSSRDIEAKFVLDKTEKNYQWLYSSASPAFQNKWSLFGYYRYLLGLDPKQT 688  
  
RESULT 14  
ADH47759  
ID ADH47759 standard; protein; 4952 AA.  
XX  
AC ADH47759;  
XX  
DT 25-MAR-2004 (first entry)  
XX  
DE NOV18 protein, SEQ ID 56.  
XX  
KW Antidiabetic; anorectic; cardiac; hypotensive; antiarteriosclerotic;  
KW anorectic; viricide; antibacterial; fungicide; protozoacide; neurotropic;  
KW neuroprotective; antiparkinsonian; anticonvulsant; osteopathic;  
KW antiarthritic; antiinflammatory; dermatological; antiasthmatic;  
KW antilipemic; Gene therapy; human; metabolic disorder; diabetes; obesity;  
KW viral infection; bacterial infection; fungal infection;  
KW helminthic infection; protozoal infection; anorexia; cancer;  
KW cardiovascular disease; neurodegenerative disorder; Alzheimer's disease;  
KW Parkinson's disease; epilepsy; immune disorder; haematopoietic disorder;  
KW inflammatory skin disorder; asthma; dyslipidaemia; NOV18;  
KW ARL-like protein; chromosome 12q12-q14.  
XX  
OS Homo sapiens.  
XX  
XX WO200268647-A2.  
XX  
XX 06-SEP-2002.  
XX  
XX 16-JAN-2002; 2002WO-US001311.  
XX  
XX 16-JAN-2001; 2001US-0261376P.  
PR 18-JAN-2001; 2001US-0262454P.  
PR 18-JAN-2001; 2001US-0262587P.  
PR 31-JAN-2001; 2001US-0265530P.  
PR 14-FEB-2001; 2001US-0268595P.  
PR 28-FEB-2001; 2001US-0272409P.  
PR 16-MAR-2001; 2001US-0276777P.  
PR 17-MAY-2001; 2001US-0291672P.  
PR 27-SEP-2001; 2001US-0325306P.  
PR 18-OCT-2001; 2001US-0330336P.  
PR 09-NOV-2001; 2001US-0345202P.  
XX  
XX (CURA-) CURAGEN CORP.  
XX  
XX Padigar M, Alsobrook JP, Colman SD, Spytek KA, Boldog F;  
XX Vernet CAM, Li L, Shenoy S, Casman S, Guo X, Edinger S;  
XX McDougall J, Malyankar U, Patturajan M, Shinkets RA, Pena C;  
XX Tchernev V, Zernuhen BD, Millett I, Miller C, Lepley DM, Smithson G;  
XX Baumgartner J, Herrmann J, Peyman JA, Gorman L, Mezes P, Kekuda R;  
XX Taupier RJ, Gerlach V, Grosse WM, Liu X, Ellerman K, Rothenberg M;  
XX Stone DJ, Burgess CE;  
XX  
XX WPI; 2002-698671/75.  
DR N-PSDB; ADH47758.  
XX  
XX New isolated NOX polypeptides and polynucleotides, useful for  
XX preventing, diagnosing or treating NOX-associated disorders e.g.  
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,  
PT asthma, or infections.  
XX  
XX Claim 1; Page 159-160; 380pp; English.  
XX  
XX The present invention relates to novel proteins (I) referred to as NOXV,  
XX where x is any number from 1 to 18, and their coding sequences (II). The  
CC proteins and their coding sequences are useful in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease,  
CC preferably a NOXV-associated disorder such as metabolic disorders,  
CC diabetes, obesity, infectious diseases (viral, bacterial, fungal,  
CC helminthic, and protozoal), anorexia, cancer, cardiovascular diseases  
CC (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:43:10 ; Search time 14.4085 Seconds  
(without alignments)  
2784.626 Million cell updates/sec

Title: US-09-155-676B-5

Perfect score: 2190

Sequence: 1 IGVTWRRRSRIVDPRAAXNS.....RGLESSUSIPSTGRGRTHF 417

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134	6.1	423	2 T05511	hypothetical prote
2	118.5	5.4	1479	2 T17401	transcription regu
3	110.5	5.0	673	2 S25335	transcription fact
4	109	5.0	4957	2 T03455	ALR protein - huma
5	109	5.0	5262	2 T03454	ALR protein - huma
6	106.5	4.9	850	2 T13352	stn-A protein - fr
7	105	4.8	544	2 AH2494	hypothetical prote
8	103.5	4.7	753	2 S48059	metal-regulatory t
9	103.5	4.7	1980	2 S54307	myosin heavy chain
10	103.5	4.7	3224	1 S58884	Ran-binding protei
11	103.5	4.7	3498	2 T22330	hypothetical prote
12	103	4.7	3449	2 T01083	hypothetical prote
13	102.5	4.7	1009	2 S60248	protein-tyrosine k
14	102	4.7	762	2 B86373	protein T23823.10
15	101	4.6	359	2 T01275	hypothetical prote
16	101	4.6	467	1 A49377	involucrin - mouse
17	101	4.6	560	2 S54037	hypothetical prote
18	101	4.6	2101	2 A42184	nuclear mitotic ap
19	100.5	4.6	771	1 A33430	h-caldesmon - chic
20	100.5	4.6	1277	2 S53043	probable membrane
21	100	4.6	810	2 G01252	small GTP binding
22	100	4.6	1027	2 T46481	hypothetical prote
23	100	4.6	1031	2 D88912	protein T06A10.1
24	100	4.6	1031	2 T33655	hypothetical prote
25	100	4.6	1106	2 T31742	hypothetical prote
26	99	4.5	475	1 A37241	52k autoantigen Ro
27	99	4.5	1805	2 A34736	nestin - rat
28	98.5	4.5	357	2 T17027	MYB-related transc
29	98.5	4.5	1009	2 A57434	protein-tyrosine k

Query Match 6.1%; Score 134; DB 2; Length 423;  
Best Local Similarity 18.2%; Pred. No. 0.031;  
Matches 73; Conservative 58; Mismatches 153; Indels 116; Gaps 13;

Qy 44 LKEAFERLLPQVEAARKAIRAAOVRVVPHERCCWLCGCEVREHLHGNTLVLYGGL 103  
Db LSSLLDRFRSKIADVRFFLNPSVLRPQEQSNRVMCVFDEDIVELGS-----SFACSKA 91  
Qy 104 LEHLASPEHKKATNKFWMENKAQVQMKERFLVTPQDYA----- 141  
Db INHFASSDHLKNIKQFLSKNGPAMDCIDFRISADVAKWEKKQSGFNEDASFEQSGQ 151  
Qy 142 -----RPFKSMVKGLDSYEKEDKVIKEMAAQIREVQSGREVRSV 183  
Db LSGTSNDIHTKLAFTMDRIKKVPAHHINSYKSN-----VMPQYNTNEYQISLSEIPGV 208  
Qy 184 -----LEPQAVPDPEEGSSA-----PRWKG-----MNSQVASSLQOPSND 220  
Db HNGSYLNMDSQPLCDSENGFGEHSIPCRSKDSYNGNYCTOENTQVQDKKQIDGSY 268  
Qy 221 LPPAPELDMMETGSLTFIIGHQDIPGVGNHSGATPPWM-IQDEEYIAGNQEIQPSYBEF 279  
Db -----VVGWTSISSSHSTTAG-GNVHSGAPPWPLDANDGDF--SSVQLNQSDVAR 319  
Qy 280 LKKEKEKQKLKLPDDRVRGANFDH-----SSRTSAGWLPSPGPRLEXWT-- 322  
Db FOAKVPCKNRKLPNRPVGAWAERRKIEIEMKSGHVTKSNIDPDPLNPFGRVQSGTRK 379  
Qy 323 -----PLAVQTSPTKLLQXRSRSHIOKKA 347  
Db ESRKEFEKEKRLVKTESISTESEPVKIQ-----PYISKRA 415

hypothetical protein F13M23.40 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C;Accession: T05511  
R;Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohelsel, J.; Mewes, H.W.; Mayer, K.F.  
submitted to the Protein Sequence Database, February 1999  
A;Reference number: Z15419  
A;Accession: T05511  
A;Molecule type: DNA  
A;Residues: 1-423 <BEV>  
A;Cross-references: UNIPROT:Q9SM36; EMBL:AL035523  
A;Experimental source: cultivar Columbia; BAC clone F13M23  
C;Genetics:  
A;Map position: 4  
A;Introns: 88/2; 130/3; 228/1; 253/3; 272/1  
A;Note: F13M23.40  
C;Superfamily: Arabidopsis thaliana hypothetical protein F13M23.40

#### ALIGNMENTS

##### RESULT 1

T05511

hypothetical protein F13M23.40 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C;Accession: T05511

R;Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hohelsel, J.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15419

A;Accession: T05511

A;Molecule type: DNA

A;Residues: 1-423 <BEV>

A;Cross-references: UNIPROT:Q9SM36; EMBL:AL035523

A;Experimental source: cultivar Columbia; BAC clone F13M23

C;Genetics:

A;Map position: 4

A;Introns: 88/2; 130/3; 228/1; 253/3; 272/1

A;Note: F13M23.40

C;Superfamily: Arabidopsis thaliana hypothetical protein F13M23.40

RESULT 2  
Tl7401  
transcription regulator WBSCR9 - mouse  
N;Alternate names: Williams-Beuren syndrome deletion transcript 9 homolog  
C;Species: Mus musculus (house mouse)  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: Tl7401  
R;Peoples, R.J.; Cisco, M.J.; Kaplan, P.; Francke, U.  
Cytogenet. Cell Genet. 82, 238-246, 1998  
A;Title: Identification of the WBSCR9 gene, encoding a novel transcriptional regulator,  
A;Reference number: Z18735; MUID:99077764; PMID:9858827  
A;Accession: Tl7401  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1479 <PEO>  
A;Cross-references: UNIPROT:Q9Z277; EMBL:AF084480; NID:G4165088; PIDN:AA000  
C;Genetics:  
A;Gene: Wbscr9  
A;Map position: 5  
F;1360-1415/Domain: bromodomain homology <BRO>

Query Match	5.4%;	Score 118.5;	DB 2;	Length 1479;
Best Local Similarity	21.1%;	Pred. No.2.2;		
Matches	95;	Conservative	49;	Mismatches 118;
				Indels 189;
				Gaps 24;
Qy	35	FTAASTORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEVREHL	SHG	94
Db	23	FTIPHTQ---EAF-RTRERYE-----ARLERY---SERIWTCKSTG-----SS	58	
Qy	95	NLTVLVGLLEHLASPHHKATNKFWMENKAQVQ--MKEKFLVTPQDYARFKKSMV---	148	
Db	59	QLT-----HKEA-----WEEQVEAELLKEEF--PNWYKLVLEWVWHNT	96	
Qy	149	-----KGLDS-----YBEKEDVKIKEMAAQITREVEQSRQEVSVLEPQA	188	
Db	97	ASLEKLVDSAWLEIMTKYAVGEECDPEVGKEKMLVKVIKVIHPLEK-----VDEEA	147	
Qy	189	VPDPEEGS-SAPRSWKGNQSVASSLQQ-----PSNLD--	220	
Db	148	VEKKSADGCDSPSDKENSQMAQDLQKETVVKEDGRRRESINDRARRSPRKLPTSLKK	207	
Qy	221	-----LP-----PAPELDWMETGPSLT---FICHQDIPGVGN	249	
Db	208	GERKWAPEKFLPHKYDVKLQNEDKIISNVFADSLIRTERPPNKEILYFIHKNAL---	263	
Qy	250	IHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEK-----OKLKKLPDDRAGVGNF	301	
Db	264	AGTGENAPWVEDE--LVKKYISLPKSDFDLLDPKYKWTINLPSTKRRNTGSPDRKPSK--	319	
Qy	302	HSRSTASGWLPSFGPRLEXTWPTLAVQSTSTPKLKLQXRSS-----	341	
Db	320	KPKRDSSSLSPNLNPKLVCHVHLEKSLNGPPLKVYKNSKNKSKEEHLGVYKIMSPNNK	379	
Qy	342	-----HIQKXANHAIYQIPXGKGP--KSTKPL	366	
Db	380	LHSFPHIPKKG-----PAAKKPKGSHDKPL	403	

RESULT 3  
S35335  
transcription factor MTF-1 - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
C;Accession: S35335; S36946  
R;Radtke, F.; Heuchel, R.; Georgiev, O.; Hergersberg, M.; Gariglio, M.; Dembic, Z.; Schaub  
EMBO J. 12, 1355-1362, 1993  
A;Title: Cloned transcription factor MTF-1 activates the mouse metallothionein I promoter  
A;Reference number: S35335; MUID:93223676; PMID:8467794  
A;Accession: S35335  
A;Molecule type: mRNA  
A;Residues: 1-673 <RAD>  
A;Cross-references: UNIPROT:O9UJW6; EMBL:X71327

R;Radtko, F.  
submitted to the EMBL Data Library, May 1993  
A;Reference number: S36946  
A;Accession: S36946  
A;Molecule type: mRNA  
A;Residues: 1-61, 63-93, 'Q', 94-283, 'N', 285-673 <RA>  
A;Cross-references: EMBL:X71327  
C;Genetics:  
A;Gene: MTF-1

Query Match 5.0%; Score 110.5; DB 2; Length 673;  
Best Local Similarity 21.3%; Pred. No. 3;  
Matches 73; Conservative 47; Mismatches 134; Indels 89; Gaps 15;

Qy 21 IRAHEG--PPFC---GRGTFRTASTORQLKEAFERLLPQVEAARKAIRAAQVERYYPEH- 74  
Db 249 IRTHTEKFPKPRDHDCGKAFRAS-----HHLKTHVRVTT 283  
Qy 75 -ERCWCCLCGCE-----VREHL-SHGNLTVLGYGLLEHLASPEHKK----- 114  
Db 284 GERPFPCPSNGCEKTFSTQVSLKSHMKGHDNKGATATYALPHNGSEDYTNHSLVLSLGLL 343  
Qy 115 ATKFWFNKNAEVOMKEKFLVTPQDYARFKGMVGLDSYEKEKDKVIKEMAAQIREVEQ 174  
Db 344 STDSELQENSSSTQDQLSTISP---AIFESM-----FQNSDDPGIQDDPLQTAALID 394  
Qy 175 SRQEVVRVLEPOAVDPDEEGSSAPRSWKGMSQVASSLQQPSNLPLPAPELDWMETGP 234  
Db 395 SFNGDAESVID---VPPP-----AGNSASLSPLVLQSGISEPPQPLLP--ATAP 439  
Qy 235 SLTFIGHQDIPGVGNTHSGA-----TPWMVQDEEY- IAGNOEIGPSYEEFLKEKEKQK 287  
Db 440 SAP-----PPAPSLPGSQAPAFGSPALLQPPVPVPVPHSTQFAANHQEFPLPHQAPP 492  
Qy 288 LKKLPPDRVCANFDHSRSISAGWLPFGFRLEXTWTLAVQTST 330  
Db 493 QTIVPGLSVVAGAPASAAATVASA AAPPPQSTTEPLPAWVQT 535

RESULT 4  
T03455  
ALR protein - human  
C;Species: Homo sapiens (man)  
C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C;Accession: T03455  
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, Oncogene 15, 549-560, 1997  
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homology to the human MTF-1  
A;Reference number: Z14954; MUID:97388474; PMID:9247308  
A;Accession: T03455  
A;Status: preliminary;  
A;Molecule type: mRNA  
A;Residues: 1-4957 <PRA>  
A;Cross-references: UNIPROT:O14686; EMBL:AF010404; NID:G2358286; PIDN:AACS1735.1; PID:97388474  
C;Genetics:  
A;Gene: ALR  
A;Map position: 12  
C;Superfamily: acute lymphoblastic leukemia protein, ALR type  
C;Keywords: alternative splicing

	Query Match	5.0%	Score 109;	DB 2;	Length 4957;
	Best Local Similarity	20.9%;	Pred. No. 55;		
	Matches	80;	Conservative 51;	Mismatches 149;	Indels 102; Gaps 15;
Qy	101	GGLEHLHASP-	-----HKATNKFWENKA-----	-----EVQMKXFLVTPQDYAR	142
Db	2609	GGPFAHLTSP	SPSGSSLLKFKFELSGALTLPGGPAA	SGDELDMXESSLVASE-----	2664
Qy	143	FKKSMVGLD	SYBEKDKVTKENAAQIREVEQSRQEV	VRSVLSP-----QAVDPPEGGSS	197
Db	2665	-LPLLIEDL	LEHEBKKLQKKQQAQLOPAAQQQQQ	QQQSHSLPAPGPAQMSLPHEGSS	2723
Qy	198	APRSWTKMNSQ	VA-----SSILQOPSNLDLPAP	ELDMWTGSPSLTFIGHQ-----	242

Db 2724 P--SLAGSQQLSLGLAVARQGLPQPLMPPTQPPAHALQ--QRLAPSMAMVSNQGHMLSGQ 2780  
Qy 243 -----DIPGVGNIHSGATPPMMIODEEVIAGNQIGPSY-----EEF 279  
Db 2781 HGGAGLVPOSSQPVLSQKPMGTWPPSCMKPQOLAMQOOLANSFFPDDTDLCKFAEDI 2840  
Qy 280 LKEKEKOK-----LKKLPDPVRGANFDHSSRTSAGWLPSPGPRLEKWTPLAVTSTPKL 333  
Db 2841 IGPIAKAKWALKGKIKV-----MAQGSIGVAPGNMRQ-----QVSLLAQ 2880  
Qy 334 KLOQXRSHIQKANKHALYQLPXGKPKSTKPLLYLPKPF--IIVLRKQTSYFIKFN- 390  
Db 2881 RLSCGSPSDDLQNHVAAGSQRSAGDSQPRP---NPPTFAQGVINEADQRQYEWLFT 2937  
Qy 391 -KVLFFGLRGLESLSIPSTSR 411  
Db 2938 QQLQMLQKLVLEEQIGVHRKSR 2959  
RESULT 5  
T03454  
ALR protein - human  
C;Species: Homo sapiens (man)  
C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C;Accession: T03454  
R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, Oncogene 15, 549-560, 1997  
A;Title: Structure and expression pattern of human ALR, a novel gene with strong homolog A;Reference number: Z14954; MUID:97388474; PMID:9247308  
A;Accession: T03454  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-5262 <PRA>  
A;Cross-references: UNIPROT:O14686; EMBL:AF010403; NID:92358284; PIDN:AAC51734.1; PID:92 C;Genetics:  
A;Gene: ALR  
A;Map position: 12  
A;Superfamily: acute lymphoblastic leukemia protein, ALR type  
C;Keywords: alternative splicing  
Query Match 5.0%; Score 109; DB 2; Length 5262;  
Best Local Similarity 20.9%; Pred. No. 60;  
Matches 80; Conservative 51; Mismatches 149; Indels 102; Gaps 15;  
Qy 101 GGLLEHLASPE-----HKATNKFWMENKA-----EVQMKKFLVTPQDYAR 142  
Db 2914 GPPPAHLTPPLSGPGGSSLLERFELESGLTLPGGPAASGDELDKXWESSLVASE----- 2969  
Qy 143 FKSMVKGSLDSYEKEDKVIKEMAAQIREVRSQREVRSVLEP-----QAVPDPERGSS 197  
Db 2970 -LPLIEDLLEHEKELQKQLSAQLOPAQOQQOQQOQHSLLPAPGAQAMSLPHEGSS 3028  
Qy 198 APRSNKGMNSQVA-----SSIQPSNLDLPAPPELDMMETGSLTFTGHQ----- 242  
Db 3029 P--SLAGSQQLSLGLAVARQGLPQPLMPPTQPPAHALQ--QRLAPSMAMVSNQGHMLSGQ 3085  
Qy 243 -----DIPGVGNIHSGATPPMMIODEEVIAGNQIGPSY-----EEF 279  
Db 3086 HGGAGLVPOSSQPVLSQKPMGTWPPSCMKPQOLAMQOOLANSFFPDDTDLCKFAEDI 3145  
Qy 280 LKEKEKOK-----LKKLPDPVRGANFDHSSRTSAGWLPSPGPRLEKWTPLAVTSTPKL 333  
Db 3146 IGPIAKAKWALKGKIKV-----MAQGSIGVAPGNMRQ-----QVSLLAQ 3185  
Qy 334 KLOQXRSHIQKANKHALYQLPXGKPKSTKPLLYLPKPF--IIVLRKQTSYFIKFN- 390  
Db 3186 RLSCGSPSDDLQNHVAAGSQRSAGDSQPRP---NPPTFAQGVINEADQRQYEWLFT 3242  
Qy 391 -KVLFFGLRGLESLSIPSTSR 411  
Db 3243 QQLQMLQKLVLEEQIGVHRKSR 3264

RESULT 6

T13352

stn-A protein - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T13352

R;Kelly, L.

submitted to the EMBL Data Library, May 1998

A;Reference number: Z17660

A;Accession: T13352

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-850 <XEL>

A;Cross-references: UNIPROT:Q24211; EMBL:U54982; NID:93138877; PID:93138878; PIDN:AAC166

C;Genetics:

A;Cross-references: FlyBase:FBgn0016976

Query Match 4.9%; Score 106.5; DB 2; Length 850;

Best Local Similarity 23.5%; Pred. No. 8.1;

Matches 90; Conservative 52; Mismatches 150; Indels 91; Gaps 22;

Qy 38 ASTORQLKEAFERLL--POVEAA--RKAIRAAQVERVYVPEHERCCWCCLCCGCEVREHLSHG 94

Db 270 ANPKRLKURKIGIQLLLSEIELADSEALLAATSNABPQHNLL-----DDLDEELSES 322

Qy 95 NLTVLYGGLLEHLASPEHKATNKFWMENKAQVQMKFLVTPQDYARFKKSMVKGSLDSY 154

Db 323 SVPIDLSVSL-HLHLIKHKQPVEE---EEBLEQKQRENQLLNP-DLSEF-----DSL 369

Qy 155 EEKEDKVIKEMAAQIREVRSQREVRSVLEPQAVPDPERGSSAPRSKGMNSQVASSLQ 214

Db 370 KDEEDDEFABLAAB--SLTKKEEVTWVSQVLPVA-QLPTEAFEA--GSWAEEFEQ---SGQ 423

Qy 215 QPSNLDLPAPPELDMMETGSLTFTGHQDIPGVGNIHSGATPPMMIODEEVIAGNQIEGP 274

Db 424 EPCKPKRPPPPVRP--PTGPHI-----YVSEEEENPDDPLNT 465

Qy 275 SYEEFLKEKEKQKLKLPDPVRGANFDHSSRTSAGWLPSPF-GPRLEXWTPLAVQTS--- 329

Db 466 NYAE-----QVKKTTVLEEDDDFPRAEEHATEPFLAAPQDILLAGSATDLSQVVP 518

Qy 330 ---TPKLKLOQXRSSH--IQKANKHALYQLPXGKPKSTKPLLYLPKPFIIIVLRKQTY 384

Db 519 APLAPTLSDVQAEADFDPFDTSVAVSLVQ-----PKSTE-----LRFLERELL 561

Qy 385 SFTXFNKVLFFGLRG--LESSLS 405

Db 562 N-----YSGLDGVTLKHLSL 576

RESULT 7

AH2494

hypothetical protein all7136 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120a1

C;Species: *Nostoc* sp. PCC 7120

A;Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004

C;Accession: AH2494

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, M.; DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AH2494

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-544 <KUR>

A;Cross-references: UNIPROT:Q8YL02; GB:BA000020; PIDN:BA078220.1; PID:G17135674; GSPDB:G17135674

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all7136

A;Genome: plasmid



[illegible]

```

QY 40 -----TORQLKEAFERLL-POVEAARKAIRAAQVRYV--PEH-----ERCWCCL- 81
Db 2658 IKHVEESEKELKDDGEYLIRPYLEHLEKIRFRYNRCERVVGLDKHGDIFLIGELCLYVIE 2717
QY 82 -----CCGCEVREHLS-----HGNLTUVLYGGLLEHLASPEHKKATNKPW 120
Db 2718 NFYIDDHGCIKECEDELSIIDQAQGLKKQFHGSL-----ESKSKSSTLW 2763
QY 121 -----WENKAQVQMEKFLVT---PODYAREFKKSMVKGLDSYEE--KEDKVI 162
Db 2764 STTIKIGAVGGRAWAYGGGAWGKEKRVVTGNLPHPHWMK-----LDSVHEILKRDYEL 2817
QY 163 KEMAAQIREVE-----QSRQEVRSVL-----EPOAVDPPEGS-- 196
Db 2818 RRVAVEIFSMDCNDLLVPHKKEREVEFRNLLAMNLPNSMLDTTIGSQAQESKGSRL 2877
QY 197 -----SAPRSWKGMNSQVASSIQPSNL-----DLPAPELDWM-----ETGPSLTF 238
Db 2878 FKLMAKSFTKRWO--NGEI-SNFQYLMHLNTLAGRGYSDLTQYPPVPMILADYDGESL-- 2932
QY 239 IGHQDIPGVGNHSGATPPWMIQDEYIAGNOEIGFSYEEFLKEKEKQ----- 286
Db 2933 -----DLSDDNNFRK-LDKPMGQCTPE---GEEFRRKRYESWDDPEVPQPHYGSHYSSAGI 2984
QY 287 ----KLKKLPPDRV-----GANFDHSSRTSAGWLPFGPRLEXWTPPLAVQTSPTKLK 334
Db 2985 VLFYLIRLPFSAENQKLOGGQFDHADRL-----FNSIRETLWSAAGKGNVSDVK 3034

RESULT 13
S60248
protein-tyrosine kinase (EC 2.7.1.112) CAK-beta - human
N:Alternate names: cell adhesion kinase-beta
C:Species: Homo sapiens (man)
C>Date: 10-Apr-1996 #sequence revision 27-Feb-1997 #text_change 16-Aug-2004
C:Accession: S60248; G02330; B57434
R:Lev, S.; Moreno, H.; Martinez, R.; Canoll, P.; Peles, E.; Musacchio, J.M.; Plozman, G.;
Nature 376, 737-745, 1995
A:Title: Protein tyrosine kinase PYK2 involved in Ca(2+)-induced regulation of ion chan
A:Reference number: S60248; MUID:95379967; PMID:7544443
A:Accession: S60248
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1009 <LEV>
A:Cross-references: UNIPROT:Q14289; EMBL:U33284; NID:g988304; PIDN:AAC50203.1; PID:g9883
R:Sasaki, T.; Nagura, K.; Sasaki, H.
submitted to the EMBL Data Library, December 1995
A:Reference number: H01067
A:Accession: G02330
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-22,'G',24-434,'L',436-1009 <SAS>
A:Cross-references: EMBL:U43522; NID:g1165218; PIDN:AAC05330.1; PID:g1165219
R:Sasaki, H.; Nagura, K.; Ishino, M.; Tobitoka, H.; Kotani, K.; Sasaki, T.
J. Biol. Chem. 270, 21206-21219, 1995
A:Title: Cloning and characterization of cell adhesion kinase beta, a novel protein-tyro
A:Reference number: A57434; MUID:95403356; PMID:7673154
A:Accession: B57434
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-22,'G',24-150 <SA2>
A:Cross-references: GB:D45853; NID:g1000676; PIDN:BAA08289.1; PID:gi000677
C:Genetics:
A:Gene: CAKBeta
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F:423-686/Domain: protein kinase homology <KIN>
F:431-439/Region: protein kinase ATP-binding motif

Query Match 4.7%; Score 102.5; DB.2; Length 1009;
Best Local Similarity 22.4%; Pred. No. 20;
Matches 84; Conservative 47; Mismatches 97; Indels 147; Gaps 23;

```

```

QY 119 FWMENKAEVQMEK-----FLVT-----PODYARFKKSMVKGLDSYEEK 157
Db 627 FWLEKNDVIGLEKGDRLPKPDLCPVLYTLMTRCWDYDSDRPFRFTSLVCSLSVDYQME 686
QY 158 EDKVIKEMAAQIREVQSRQEVRS--VLEPQAVDPPEEGSSAPRSWKGMNSQVASSLQQ 215
Db 687 KDIA-----MEQERNARYRTPKILEFTAFQEPKPSRPK-----YRP 724
QY 216 PSNLDLPPAPELDWM-----ETGPSLTFIGHQDIPG-VGNIHSGATPP-----WM 259
Db 725 PPQTNL-LAPKLQFQVPEGLCASSPTLT--SPMEYPSFVNSLH--TPPLHRHNVPKRS 778
QY 260 IQDEYI--AGNOEIGPSYE-----EFLKEKEKOKL-----KLLPPDRVGANF 300
Db 779 MREEDFIQPSREAAQWLWEAEKVMRQLDKQKQWVEDYQWLROEKSLLDP-MVYWN- 836
QY 301 DHSRT---SAGWLPFGPRLEXWTPPLAVQTSPTKLKLQXRSRSHIQKAN----- 348
Db 837 DKSLPTPEKEVYLEFTGP-----PQKPRLGAQS-----IOPTANLDRDLDLVY 881
QY 349 -----HALYQLPXGXKPKSTKPLLYLPPKFFIIVFLRKQTYSFIXNKVLPFLGLRGLE 401
Db 882 LNVMLVRAVLLEL-----KNELCQLPPEGYVVV-----KNVGLTLRKL 921
QY 402 SSL-----SIPSTGR 411
Db 922 GSVDDLPLPSLSSSR 936

RESULT 14
E86373
protein T23E23.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86373
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86373
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-762 <STO>
A:Cross-references: UNIPROT:Q9LRA0; GB:A5005172; NID:g9369408; PIDN:AAF87156.1; GSPDB:GN
C:Genetics:
A:Gene: T23E23.10
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein F18N11.140

Query Match 4.7%; Score 102; DB.2; Length 762;
Best Local Similarity 19.5%; Pred. No. 15;
Matches 66; Conservative 49; Mismatches 99; Indels 124; Gaps 15;

QY 118 KFWMENKAEVQMEKFLVTQD-----YARFKKSMVKGLDSYEEK---BDKV 161
Db 456 KIFWRSLYQADIDGRRLETVDDVPVPEPEIPGTSTLSFKEMERGFDKLTDKLAVWDSE 515
QY 162 IKEMAAQIREVQSRQEVRSVLPEQAVDPPE-----GSSA---PRSWKGMNSQVASSLQ 214
Db 516 IKRCIVRVQGIE-----EYVADQLEKEAEKQYEDMRFGSAADFVPPSYQPRDTQSKALV- 570
QY 215 QPSNLDLPPAPELDWMETGSLTFIGHQDIPGVGNHSGATPPMIODEYIAGNQIGP 274
Db 571 -----IHSGRDPAPTTIESLSDGSDS--- 592
QY 275 SYEEFLKEKEKQKLKLLPPDRVGANFD-----HSSRT 306

```



Db 593 SDEEDKKKKKKQISKKRKKVGVFTLESSEGRKGRKPSKYKGEBFTTECKQTRKKQKSSS 652  
Qy 307 SAGMLP-SFGPRLEXKWTPLAVQTSTPKLKLOQXRSSHQKKANHALYQL----- 354  
Db 653 SSSYRPQSSSARLQA-LPITATISSPSCQL-----VSSSDSSSSYQLIPQLQNSSTSS 704  
Qy 355 PXGKKPKSTKPLLYL---PPKFFIIVLRKQTYSFIXF 389  
Db 705 PSGS--SSIDCIVHLSRRPARLYL-----QOTTRFVQF 735

RESULT 15  
T01275  
hypothetical protein At2g19270 [imported] - Arabidopsis thaliana  
N:Alternate names: hypothetical protein F27F23.7  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T01275; F84574  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, May 1998  
A:Description: Arabidopsis thaliana chromosome II BAC F27F23 genomic sequence.  
A:Reference number: Z14177  
A:Accession: T01275  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-359 <ROU>  
A:Cross-references: UNIPROT:O64560; EMBL:AC003058; NID:g3135250; PID:g3135257  
A:Experimental source: cultivar Columbia  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: F84574  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-359 <STO>  
A:Cross-references: GB:AE002093; NID:g3135257; PIDN:AAC16457.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: F27F23.7; At2g19270  
A:Map position: 2  
A:Introns: 305/3

Query Match 4.6%; Score 101; DB 2; Length 359;  
Best Local Similarity 23.8%; Pred. No. 6.6;  
Matches 50; Conservative 30; Mismatches 84; Indels 46; Gaps 8;  
Qy 146 SMVKGLDSYEKEDKVIKEMAAQIREVQSQREVRSVLEPQAVDPPEGSAPRSWKGM 205  
Db 3 SLVASYASSDEEDLQPRREAFTVKSGEPSASSLFSALPQPKELKSSDGLGSSSRGK 62  
Qy 206 NSQVASSLQOP-SNLDLPPAPELDWMETGPSLTFIGHDIP-GVGNHSGATP-PWMIQD 262  
Db 63 SSSFLSSLPKPKSSISRKNP-----NPS-----SIPRVVQIRLPVNPSPSLDD 108  
Qy 263 EETIAGNQEIGPSYEEFLKEKEKOKLKLPPDRVGANFDHSRT-----SAGW 310  
Db 109 EDD-----EEEKARKKQKQMESASASHDSSVRSFLSAMPAPKSSQTIGA 153  
Qy 311 LPSPGPRLEXKWTPLAVQTSTPKLKLOQXRS 340  
Db 154 LPSLGGSGGRSRL--ETETPSIAFPQDTS 181

Search completed: June 7, 2005, 12:03:29  
Job time : 17.7419 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:40 ; Search time 61.0244 Seconds  
(without alignments)  
3499.208 Million cell updates/sec

Title: US-09-155-676B-5

Perfect score: 2190

Sequence: 1 IGVTWRRSRIVDPRAXNS.....RGLESSLSIPSTSGRGRTHF 417

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1498	68.4	332	2 Q86UT8	Q86ut8 homo sapien
2	1302	59.5	338	2 Q7TNQ4	Q7tnq4 mus musculus
3	747	34.1	194	2 Q8C702	Q8c702 mus musculus
4	134	6.1	423	2 Q8W36	Q8w36 arabidopsis
5	118.5	5.4	1479	1 BAI8_MOUSE	Q22277 mus musculus
6	115.5	5.3	777	2 Q6GQ06	Q6gq06 xenopus lae
7	113	5.2	496	2 Q7ZUY4	Q7zuy4 brachydanio
8	112.5	5.1	487	2 Q8B6K0	Q8b6k0 soybean dwa
9	112.5	5.1	489	2 Q91QJ6	Q91qp6 soybean dwa
10	112.5	5.1	489	2 Q8B6M4	Q8b6m4 soybean dwa
11	112	5.1	1893	2 Q8CJ14	Q8cj14 rattus norv
12	111.5	5.1	850	1 STNA_DROME	Q24211 drosophila
13	111	5.1	536	2 Q8IH09	Q8lhd9 drosophila
14	111	5.1	539	2 Q8ML54	Q8mls4 drosophila
15	111	5.1	1266	2 Q7PCX6	Q7pcx6 macaca fasc
16	110.5	5.0	464	2 Q7QU17	Q7qui7 giardia lam
17	110.5	5.0	674	2 Q80W63	Q80w63 mus musculus
18	110.5	5.0	675	1 MTJF1_MOUSE	Q7243 mus musculus
19	110.5	5.0	675	2 Q9JJW8	Q9jjw8 mus musculus
20	110.5	5.0	1878	2 Q71A35	Q71a35 mycoplasma
21	110	5.0	496	2 Q8P0V1	Q8p0v1 brachydanio
22	110	5.0	675	2 Q8BSY2	Q8bsy2 mus musculus
23	109.5	5.0	446	2 Q7XT92	Q7xt92 oryza sativ
24	109.5	5.0	489	2 Q8B6K2	Q8b6k2 soybean dwa
25	109.5	5.0	850	2 Q8XQ55	Q8xqes raietonia s
26	109.5	5.0	1109	2 Q94HE0	Q94he0 oryza sativ
27	109.5	5.0	1109	2 Q7XFH5	Q7xfh5 oryza sativ
28	109	5.0	561	2 Q7XWH8	Q7xwh8 oryza sativ
29	109	5.0	1981	2 Q923Z7	Q923z7 rattus norv
30	109	5.0	5262	1 MLL2_HUMAN	O14686 homo sapien
31	108.5	5.0	1225	1 CTD2_HUMAN	Q9uqb3 homo sapien

32 108.5 5.0 1984 2 Q7TNI3  
33 108 4.9 844 2 Q8L220  
34 108 4.9 844 2 Q7DJM2  
35 108 4.9 1255 2 Q86167  
36 108 4.9 1255 2 Q79VT6  
37 107.5 4.9 999 2 Q86UE5  
38 107.5 4.9 1454 2 Q7ZWK4  
39 107.5 4.9 1871 2 Q719M1  
40 106.5 4.9 562 2 Q6A2A5  
41 106.5 4.9 3053 2 Q9ERU9  
42 106 4.8 270 2 Q9FR84  
43 106 4.8 385 2 Q54913  
44 106 4.8 476 2 Q6GR02  
45 106 4.8 612 2 Q8BL65

#### ALIGNMENTS

RESULT 1  
Q86UT8 PRELIMINARY; PRT; 332 AA.  
AC Q86UT8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE DLNB14 protein.  
GN Name=DLNB14;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kubo T., Arai Y., Ohira M., Gamou T., Maeno G., Sakiyama T.,  
RA Toyoda A., Hattori M., Sakaki Y., Nakagawara A., Ohki M.,  
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB094093; BAC76047.1; -  
SQ SEQUENCE 332 AA; 37974 MW; 341209B75BE56349 CRC64;

Query Match 68.4%; Score 1498; DB 2; Length 332;  
Best Local Similarity 89.1%; Pred. No. 4.3e-95;  
Matches 286; Conservative 5; Mismatches 24; Indels 6; Gaps 2;  
Qy 28 FFCRGFTTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87  
Db 14 FFCRGHVYRKQRQKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73  
Qy 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAQVQMEKFLVTPQDYARFKKSM 147  
Db 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAQVQMEKFLVTPQDYARFKKSM 133  
Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVPPPEGSSAPRSWKGNMS 207  
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVPPPEGSSAPRSWKGNMS 193  
Qy 208 QVASSLQOQSNLDPAPPELDMMETGSLTFIGHQDIPGVGNTHSGATPPMMIODEEVIA 267  
Db 194 QVASSLQOQSNLDPAPPELDMMETGSLTFIGHQDIPGVGNTHSGATPPMMIODEEVIA 253  
Qy 268 GNOEIGPSYEEFLKEKEKQKLLKLPDRVGANFDHSSRTSAGWLPSPFGRLXWTPLAVQ 327  
Db 254 GNOEIGPSYEEFLKEKEKQKLLKLPDRVGANFDHSSRTSAGWLPSPFGRLXWTPLAVQ 310  
Qy 328 TSTPKLKLQ---QXRSSHIQK 345  
Db 311 QSRHQFKEAAAMKQSHTEK 331  
RESULT 2  
Q7TNQ4 PRELIMINARY; PRT; 338 AA.  
ID Q7TNQ4  
AC Q7TNQ4;

```
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE D63004F24Rik protein (Fragment).
GN Name=D63004F24Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon; DOI=10.1073/pnas.242603899;
RX MEDLINE=2238257; PubMed=12477932;
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Colon;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055929; AAH55929.1;
FT NON_TER
SQ SEQUENCE 338 AA; 38244 MW; 8F0858468908CBA1 CRC64;

Query Match 59.5%; Score 1302; DB 2; Length 338;
Best Local Similarity 78.1%; Pred. No. 1.4e-81;
Matches 246; Conservative 19; Mismatches 40; Indels 10; Gaps 1;

QY 28 PFCGRGTTAATQQLKEAFRLLPQVEAARKATRAAQVERYPVEHRCWCWCCGCEV 87
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
20 PFCGRGHVYSHKHQRLKALERLLPQVEAARRAVRAAQVERYPVEHRCWCWCCGCEV 79
QY 88 REHLHGNIIVLYGLLEHLASPEHKKATNKFWMENKAEVOMKEXFLYTPDYARFKSM 147
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
80 RKHLHGNIIVLYGLLEHLASPEHKKATNKFWMENKANAQMKEXFLYTPDYARFKSM 139
QY 148 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRVLPQAVPDPDEGSSAPRSWKGNIS 207
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
140 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRVLPQAVPDPDEGSSAPRSWKATNG 199
QY 208 QVASSLQPSNLDLPAPFELDMWETGTPSLFTFGHQDIPGVGNHSGATPPWMIQDEYIA 267
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
200 HVASSSQVSHLALQFVAELDMWETGTPSLFTFGHQDIPGVGNHSGATPPWMIQDEYIS 259
QY 268 NQGETGPSYEELKEKEKOKLKLPPDRVGANFDSHSSRTSAGWLPSPG-----PR 317
Db ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
260 GSLPFGPSYEELKEKEKOKLKLPPDRVGANFDSHSSNTSAGWLPSPGFRVWNNRRWQSR 319
QY 318 LEXWTPLAVOTSTPK 332
Db : : : :
320 HQFKTEATRSPQR 334

RESULT 3
Q8C702
```

```
ID Q8C702 PRELIMINARY; PRT; 194 AA.
AC Q8C702;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
DE library, clone:D63004F24 product:hypothetical protein, full insert
DE sequence.
GN Name=D63004F24Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno N., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
```

Birney E., Hayashizaki Y.;  
"Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."; *Nature* 420:563-573(2002).



```

RL Dev. Dtn. 225:384-391 (2002).
RP [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Gerhard D.S.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC072944; AAH72944.1; -.
DR InterPro; IPR004022; DDT.
DR SMART; SM00571; DDT; 1.
DR PROSITE; PS0827; DDT; 1.
FT NON TER 777
SQ SEQUENCE 777 AA; 89051 MW; 8DBABF7365DE2087 CRC64;

Query Match 5.3%; Score 115.5; DB 2; Length 777;
Best Local Similarity 21.0%; Pred. No. 21;
Matches 100; Conservative 55; Mismatches 151; Indels 171; Gaps 24;

Qy 31 GRGTFTAASTORQKEAF---ERLLPQVEAARAKIRA---AQVERVYFEHERCCWCL 81
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6 GRPFPPLV---KPLSEAAATGGEVEVYMIHSKFAFRSEYERLRYA---ERITWCK 59
Qy 82 CCGC-----EVRHLSHGNTLVLY-GGLLE--HLASPEHKKATNKPWENKA 125
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
60 STGSQSLTHKEAWDEEQEVAELLKE-EPPVWYEQVLEWHVHTISLDKLVDSQWMEIMT 118
Qy 126 EVQMKKE--FLVTPQDYARFKKSMVKGGLDSYE-----EKEDKVIKEM 165
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
119 KYADGEECDFFVGPVKYLRAKIVKHPLEKEQASEKSEKSGCDSPSSDKENSNAQDI 178
Qy 166 AAQREVESQREVRSVLE-----PQAVP---DPEGSSAPRSWKGMNSQVASSLOQ 215
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179 --QLKE-ESNLESLSSRESDRARRSPKLTSLKKEKKWPPKFLPHKYDV-KLINE 234
Qy 216 PSNLDLPAPELDWMETGSLT---FICHQDIPGVGNHSGATPPWMIQDEEYIAGNQE 271
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
235 DKVISFVPVDSLYSRPPNKEILRYFIRNAL-----RIGTENAPWVEDE--LVKKYT 288
Qy 272 IGPSYERFL-----KEKEKOKLKLPPD----- 294
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
289 LPSKFSDFLLPHKYMTLNPSSATKRKSLGSPDQKPAKSKKSLSPSSWSLANLKKTAV 348
Qy 295 -----RVGANFPH-----SSRTSAGWLPSPG-- 315
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
349 NSSSESEEMQLMGLANLNKSGSIGKSKDKPKNGKSOVLNGQKISAKTRS---PKGLK 405
Qy 316 -PRLEXWTPLAQVTSPLKLVQXKRSHIQQKANHALVQLPXGKPKSTKPLLYLPP 371
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
406 SPKLQMTLLDMAKSTPKVSRQAGGSNTPR-----SSSKENKYLPP 447

RESULT 7
Q7ZUY4
ID Q7ZUY4 PRELIMINARY; PRT; 496 AA.
AC Q7ZUY4;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Nucleobindin 2a.
GN Name=nuc2a;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.I., Wang J.I., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnertch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046077; AAH46077.1; -.
DR ZFIN; ZDB-GENE-030826-14; nucb2a.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR010983; EF_Hand_Like.
DR Pfam; PF00036; ehand; 2.
DR SMART; SM00054; EFh; 2.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
SQ SEQUENCE 496 AA; 58136 MW; 4FD038C12A455AE2 CRC64;

Query Match 5.2%; Score 113; DB 2; Length 496;
Best Local Similarity 20.8%; Pred. No. 17;
Matches 64; Conservative 47; Mismatches 94; Indels 102; Gaps 16;

Qy 13 DPRAAXNSIRAHGPFPCGRGTF--TAATQRLKEAFERLLPQVEAARAKIRAQAQVERVVP 72
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
246 DPKTFFNLHDTNGDGF--DEQELSLFTKLEKI-----YDP 281
Qy 73 -----EHERCCWCLCGCEVREHLSH---GNLTVLYGGILLEHLASPEHKKATNKF 119
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
282 TNEEDDWMVEEER-----LRMRHVNVNDSNKDLV-SLDEFVLATKKKEFLPDP 332
Qy 120 WVENKAQVQKKEFLVTPQDYARFKKSMV-----KGLDSYEEKED--KVIKEMAAQI 169
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
333 SWET-----LEQNAVTEEMREFEQLVQEEDLNQKAADLQKQREDLERQEQQLNAQK 387
Qy 170 REVEQSQEVRSVLEPQAVDPDE---EGSSAPRSWKGMNSQVASSLQQPSNLDLPAPPE 226
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
388 IEQQAVHEMER--IKTKVPPPPSEILGNAPVES-----VGQDQPPVP- 429
Qy 227 LDWMTGSLTFIGHQDIPGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKQ 286
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 LEHQPLP-----GHQDVP-----PPAAQHHDELQNLQALQGHQNTHENQ-- 472
Qy 287 KLKKLPP 293
Db |||
473 ---PLPP 476

RESULT 8
Q8B6K0
ID Q8B6K0 PRELIMINARY; PRT; 487 AA.
AC Q8B6K0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Readthrough protein (fragment).
GN Name=ORF5;
OS Soybean dwarf virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
OC Luteovirus.
OX NCBI_TaxID=12049;
RN [1]

```





```
QY 355 PXGK-KPKSTKPLLYLPKPKF 373
DB 454 STGTLKGLSLKPKKELPPRF 473

RESULT 11
Q8CJ14 PRELIMINARY; PRT; 1893 AA.
AC Q8CJ14;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
OS Nestin.
DE Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=22573441; PubMed=12686602;
RA Chou Y.H., Khoun S., Herrmann H., Goldman R.D.;
RT "Nestin promotes the phosphorylation-dependent disassembly of vimentin
intermediate filaments during mitosis.";
RL Mol. Biol. Cell 14:1468-1478(2003).
[2]
SEQUENCE FROM N.A.
RA Chou Y.-H., Herrmann H., Goldman R.D.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the intermediate filament family.
DR EMBL; AF538924; AAN33053.1; -.
DR HSSP; P08670; 1GK7.
DR GO; GO:0005882; C:intermediate filament; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR011000; Apolp_III_like.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
DR PROSITE; PS00226; IF; 1.
DR KW Intermediate filament.
SQ SEQUENCE 1893 AA; 208796 MW; 03AE6B616A1A7623 CRC64;

Query Match 5.1%; Score 112; DB 2; Length 1893;
Best Local Similarity 20.6%; Pred. NO. 1.1e-02;
Matches 73; Conservative 53; Mismatches 131; Indels 98; Gaps 14;

QY 35 FTAASQ--ROLKEAFERLLPQ--VEAARKAIRAAQVRYVPEHERCCWCLCCGVEYEH 90
DB 663 FPGAEQMLRLVKEKQSFPSPEEDQEAQCRPLQKENQEP-----LGVEEAE- 711
QY 91 LSHGNLTVLVYGLLEHLASPEHKATKNKFWENKAQVQMEKFLVTPQDY-ARFKKSMVK 149
DB 712 -----QILLERLIEKSQES-----LRSPEEDQEAQGRSLQK 743
QY 150 GLD---SYEKEDKVI-----KENAAQIREVEQSRQEVVRVLEPQAVDPPEEGSSAPRS 201
DB 744 QNQPPLGYEEAGQILERLIEKSQESLSRAEEDQEAQCRSL-----QKENQEPFLG 794
QY 202 WKGNWSQVASSLQPSNLDLPAPELDMWETGPSLTFIGHQDIPGVGNHSGATPPWMIQ 261
DB 795 YEEAEDQLERLIEKSQESLSRAEEDQEAQGRSLQ-----K 831
QY 262 DEVIAGNQETGPSVEEFLKEKQKLLPPD--RVGANFDHSRTSAGWLPSPGPRLE 319
DB 832 ENQEPGLYEEAEDQMLERLIEKSQESLSRAEEDQEAQGRSLYLER---NQE 888
QY 320 XWTPPLAVQTSFKPLK-----QQRSSHIQKANHLYQLPXPQKPKSTKPLLYL 369
DB 889 TFVPLESRNQRPLRSLEVEEERIVKPLKVSQDSLSGL-----AEENVQLRYL 939

RESULT 12
STNA_DROME
ID STNA_DROME STANDARD; PRT; 850 AA.

AC Q24211; Q9WSM8;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Stoned A protein (StonedA) (Stn-A).
GN Name=stnA; ORFNames=CG40306;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Oregon-R; TISSUE=CNS;
MEDLINE=97001127; PubMed=8844157;
RA Andrews J., Smith M., Merakovsky J., Coulson M., Hannan P.,
RA Kelly L.E.;
RT "The stoned locus of Drosophila melanogaster produces a dicistronic
transcript and encodes two distinct polypeptides.";
RL Genetics 143:1699-1711(1996).
[2]
SEQUENCE FROM N.A.
STRAIN=Berkely;
MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sriden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[3]
GENOME REANNOTATION.
MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tsuy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
```



GN Name=CG30412;CG30416; ORFNames=CG30416;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BT001295; AA071050.1; -  
DR FlyBase; FBgn0050416; CG30416.  
SQ SEQUENCE 536 AA; 60593 MW; C5B462B6E254C0E CRC64;  
  
Query Match 5.1%; Score 111; DB 2; Length 536;  
Best Local Similarity 29.7%; Pred. No. 26;  
Matches 33; Conservative 21; Mismatches 39; Indels 18; Gaps 4;  
  
Qy 152 DSYEEKEDKVIKEMAAQIREVEQSRQVRSVLE---PQAVDPDEEGSSAPRSKMGNSQ 208  
Db 244 ESKDAKEEIVYQYEKTERDPPDKAKVEKTIKVIKIPVDSPEAESNSPPSGIPPPSQ 303  
  
Qy 209 VASSLQOPSNL--DLPPAPELDMMETGSLTFIGHQDIPGVGNHSGATPP 257  
Db 304 LPSDIPPPSQLSDIPPPSQL-----PS-----DIPPPSQLPSGTTPP 341  
  
RESULT 14  
Q8ML54  
ID Q8ML54 PRELIMINARY; PRT; 539 AA.  
AC Q8ML54  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE CG30416-PA.  
GN ORFNames=CG30416;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Ffannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beson K.V., Benos P.V., Beremand B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Pavenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M., G.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodruff, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
melanogaster euchromatic genome sequence";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Swirskas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.;  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomics perspective";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertram B.P.,  
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
systematic review";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003450; AAM68240.2; -  
DR FlyBase; FBgn0050416; CG30416.  
SQ SEQUENCE 539 AA; 60907 MW; 03526863C861005C CRC64;  
  
Query Match 5.1%; Score 111; DB 2; Length 539;  
Best Local Similarity 29.7%; Pred. No. 27;  
Matches 33; Conservative 21; Mismatches 39; Indels 18; Gaps 4;  
  
Qy 152 DSYEEKEDKVIKEMAAQIREVEQSRQVRSVLE---PQAVDPDEEGSSAPRSKMGNSQ 208  
Db 244 ESKDAKEEIVYQYEKTERDPPDKAKVEKTIKVIKIPVDSPEAESNSPPSGIPPPSQ 303  
  
Qy 209 VASSLQOPSNL--DLPPAPELDMMETGSLTFIGHQDIPGVGNHSGATPP 257  
Db 304 LPSDIPPPSQLSDIPPPSQL-----PS-----DIPPPSQLPSGTTPP 341





; LOCATION: (388)...(388)		; OTHER INFORMATION: Xaa can be any naturally occurring amino acid	
US-09-155-676B-5			
Query Match		99.5%; Score 2178; DB 5; Length 417;	
Best Local Similarity		100.0%; Pred. No. 1.9e-155;	
Matches 417; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	IGVTRRSRIIVDPRAAXNSIRAHGPFPCGRGTTAASTORQLKEAFERLLPQVEAARK	60
Db	1	IGVTRRSRIIVDPRAAXNSIRAHGPFPCGRGTTAASTORQLKEAFERLLPQVEAARK	60
QY	61	ATRAAQERYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKF	120
Db	61	ATRAAQERYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKF	120
QY	121	WENKAQVQKEKFLVTPQDYARFKSMVKGLDSYEKEDKVIKEMAAQIREVEQSRQ	180
Db	121	WENKAQVQKEKFLVTPQDYARFKSMVKGLDSYEKEDKVIKEMAAQIREVEQSRQ	180
QY	181	RSVLPEQAVDPDEEGSSAPRSWKGMNSQVASSLQPSNLDLPPAPELDMETGPSLTIF	240
Db	181	RSVLPEQAVDPDEEGSSAPRSWKGMNSQVASSLQPSNLDLPPAPELDMETGPSLTIF	240
QY	241	HODIPGVGNHSGATPPMWIQDEEYIAGNOEIGPSYEELKEKEKQKLLPDRVGANF	300
Db	241	HODIPGVGNHSGATPPMWIQDEEYIAGNOEIGPSYEELKEKEKQKLLPDRVGANF	300
QY	301	DHSSRTSAGWLPSPGPRLEXWTPAVQTSTPKLKLQXRSASHIQQKANHALYQLPXGKKP	360
Db	301	DHSSRTSAGWLPSPGPRLEXWTPAVQTSTPKLKLQXRSASHIQQKANHALYQLPXGKKP	360
QY	361	KSTKPLLYLPKPFIIIVLRKQTSYFIENKVLFGRLGLESSLIPSTSRGGRTHF	417
Db	361	KSTKPLLYLPKPFIIIVLRKQTSYFIENKVLFGRLGLESSLIPSTSRGGRTHF	417
RESULT 2			
US-10-450-763-49939			
; Sequence 49939, Application US/10450763			
; GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc			
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES			
; FILE REFERENCE: 790CIP3/US			
; CURRENT APPLICATION NUMBER: US/10/450,763			
; CURRENT FILING DATE: 2003-06-11			
; PRIOR APPLICATION NUMBER: PCT/US01/08631			
; PRIOR FILING DATE: 2001-03-30			
; PRIOR APPLICATION NUMBER: 09/540,217			
; PRIOR FILING DATE: 2000-03-31			
; PRIOR APPLICATION NUMBER: 09/649,167			
; PRIOR FILING DATE: 2000-08-23			
; NUMBER OF SEQ ID NOS: 60736			
; SOFTWARE: Custom			
; SEQ ID NO 49939			
; LENGTH: 156			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; LOCATION: (1)...(156)			
; OTHER INFORMATION: Xaa = X or * as defined in Table 2			
US-10-450-763-49939			
Query Match		30.6%; Score 670; DB 6; Length 156;	
Best Local Similarity		77.9%; Pred. No. 8.2e-43;	
Matches 134; Conservative		4; Mismatches 12; Indels 22; Gaps 4;	
QY	6	WRRSRIIVDPRAAXNSIRAHGPFPCGRGTTAASTORQLKEAFERLLPQVEAARKAIRAA	65
Db	5	WSRRL-OPQAPA---AASGGF--GEAPAGA-----EVEAARKAIRAA	42
QY	66	QERYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKA	125

```
; Sequence 44039, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 44039
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-450-763-44039

Query Match      23.0%; Score 503; DB 6; Length 312;
Best Local Similarity 49.3%; Pred. No. 6.7e-30;
Matches 113; Conservative 14; Mismatches 38; Indels 64; Gaps 6;

Qy 40 TORQLKEAFERL-----LPQVEAARKAIRAAQVERYPVEHERCCWCLCCGCEVREHL 91
Db 108 TQPLPAAADMPAAITGITDVEAARKAIRAAQVERYPVEHERCCWCLCCGCEVREHL 167
Qy 92 SHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKKFLVTPDYARF----- 143
Db 168 SHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKKFLVTPDYARSHLYGTPMG 227
Qy 144 ----KSNVKGK-----DSYEKEDKVIKEMAAQIREVQSQRQVRSV-LEPOA 188
Db 228 SAVPSQACLEALLIIVPGACVQSEEE-----PAGSKDEPGEQVELKEEA 276
Qy 189 VPDPEEGSAPRSWKGMNSQVASSLQPPSNLDLPPAPELDMMETGPSLT 237
Db 277 EAPVEDGSQPB-----PPEPKGDATPEGEKAT 303

RESULT 5
US-10-703-032-105976
; Sequence 105976, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Cao, Yongwei
; APPLICANT: Masucci, James D.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53374)B
; CURRENT APPLICATION NUMBER: US/10/703,032
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/020,338
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 211164
; SEQ ID NO 105976
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(567)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_TA_394.pap
; US-10-703-032-105976

Query Match      5.1%; Score 124.5; DB 6; Length 567;
Best Local Similarity 23.9%; Pred. No. 0.37;
Matches 71; Conservative 46; Mismatches 99; Indels 81; Gaps 16;

Qy 35 FTAASTORQLKEAFERLLPQVEAARKAIRAAQVERYPVEHERCCWCLCCGCEVREHL 94
Db 116 FNAVERPKEREDLFENYLVELQKKEKAKAAEHEKRTAEYR----- 156
Qy 95 NLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKKFLVTPDYARFK-----S 146
Db 157 -----EFLSCDFIKA-NTQWRKVQDRLEDDER-----YARLEKIDRLDVFQD 198
Qy 147 MVKGLDSYEKEDKVIKEMAAQIREVQSQRQVRSVLEPOAVDPDPREGS-SAPRSWKGM 205
Db 199 YIRHLEKEEBEEOKIRKE---QLRRQERKRDREKQMEHEVA---DGTINAKTYWRDY 251
Qy 206 NSOVASS---LQPPSNLDLPPAPEL---DMMETGPSLTFIGHODIPGV---GNIHSGATP-- 256
Db 252 CSQIKDSRAYLAVASNLSGMPKELFDDVMEELDK-----QYDDRALIKDEKSGKIPML 307
Qy 257 -PMWI-----QDEEYI-AGNQEIGPSYE---EFLKEKEKQKLKLPDRVGANF 300
Db 308 ASWTLDFQAAVTEDEKYGKVSINIKLIYEDQIERLKEKDLKEAKK-RQRLGDNF 362

RESULT 6
US-11-097-143-32898
; Sequence 32898, Application US/11097143
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32898
; LENGTH: 850
; TYPE: PRT
; ORGANISM: DROSOPHILA
; US-11-097-143-32898

Query Match      5.1%; Score 111.5; DB 7; Length 850;
Best Local Similarity 23.2%; Pred. No. 5.8;
Matches 89; Conservative 52; Mismatches 151; Indels 91; Gaps 21;

Qy 38 ASTORQLKEAFERLL---PQVEAA-RKAIRAAQVERYPVEHERCCWCLCCGCEVREHL 94
Db 270 ANPKRLKRGIQNLLLSIESIELADSEALAAATSNAPQHNL-----DDLDEELSES 322
Qy 95 NLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKKFLVTPDYARFKSMYKGLDSY 154
Db 323 SVPIDLSVSL-HLHLIKHKQPVEE---EELEQKRGRENQLNLP-DLSEF-----DSL 369
```





	Matches	81;	Conservative	46;	Mismatches	111;	Indels	119;	Gaps	19;
	QY	31	GRGTFTAASTORQLKE---	AFERLLPQVEAARKAIRAAQVRYYPEHERCCWCLCGCEV	87					
	Dd	50	GSETETTSAILASVKELQFRLTRELEAQIV-ASQLER-							
	QY	88	REHLSHGNTLVLYGGLLHLASPHEKKATNKFWENKAEVQMKEFLVTVDYARPKSM	147						
	Dd	94	-----GSET----	GSMSSMSAE---EQFMQSQ--	115					
	QY	148	VKGLDSEYEKEDKV---	IKEMAAQIREVEOSRQEVRSVLEPOAVDPPEEGSAPRSWK	204					
	Dd	116	----DGKDIEDELTTGLELVDSICIRLSQES-----	GILDPDYSTGERPSLLSQSALQ	165					
	QY	205	MNSQVASSLOQP-----SNLDLP-----	PAPELDMETGPSL-----TFIGHQDIPG	246					
	Dd	166	LNSKPEGSFQYPASYHSNOTLALGETTSPQLPARGTQARATGQSFSGQTTSRAGHLA	GP-224						
	QY	247	VGNHSGATPPMWIQDEEVIAGNOIGI GPSYEFLKEKEKQLKLPDRVCANFDHSSRT	306						
	Dd	225	-----EPAPPPPppPREPFA-----PSLGSAFH-----	LPDAPPAAAAAALIYSST	266					
	QY	307	SAGWLPSFGPRLEXWTPLAVOTSTPKLKLOQRSSHIOKKANHAILYOLPXGXKP	KST363						
	Dd	267	----LPA-PPR--GGSLAAPOGGSPTKLORGGA-----	PEGATYAAPRGSSPFQKS	311					
		RESULT 11								
		US-10-940-774A-7037								
		; Sequence 7037,	Application US/10940774A							
		; GENERAL INFORMATION:								
		; APPLICANT:	VENTER, J. Craig et al.							
		; TITLE OF INVENTION:	POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF							
		; FILE REFERENCE:	CLO01307							
		; CURRENT APPLICATION NUMBER:	US/10/940,774A							
		; CURRENT FILING DATE:	2004-09-15							
		; PRIOR APPLICATION NUMBER:	60/241,755							
		; PRIOR FILING DATE:	2000-10-20							
		; PRIOR APPLICATION NUMBER:	60/237,768							
		; PRIOR FILING DATE:	2000-10-03							
		; PRIOR APPLICATION NUMBER:	60/231,498							
		; PRIOR FILING DATE:	2000-09-08							
		; NUMBER OF SEQ ID NOS:	207012							
		; SOFTWARE:	FastSeq for Windows Version 4.0							
		; SEQ ID NO 7037								
		; LENGTH:	1540							
		; TYPE:	PRT							
		; ORGANISM:	Human							
		US-10-940-774A-7037								
		Query Match								
		Best Local Similarity	4.9%; Score 107.5; DB 6; Length 1540;							
		Matches	104; Conservative 60; Mismatches 126; Indels 207; Gaps 30;							
		QY	14 PRAAXNSIRAHEGPPFCGCG-----TFTAASTORQL--KEAFERLLP-----OVEA	57						
		Dd	26 PGAWRS LRA -AGPSAAPRGPRQQPVPTAAAAMAPLLGRKPLVPKPLPGEEPLFIPIH	84						
		QY	58 ARKAIRA-----AQERYVYPERHCWCCLCGCVRSHLNHTVLYGGLLHLASPEH	112						
		Dd	85 TQEAAPTREYEARELRV--SERITWCKSTG-----SSQLT-----H	119						
		QY	113 KKATNKFWMENKAEVQ-MKEKFLVTVDYARPKKSMV-----KGLDS-----	153						
		Dd	120 KEA-----WESEQVAELLKEEF---PAWYEKLVMVHHNTASLEKLVDTAMLTEIMTKY	171						
		QY	154 -----YEKKEDVKIKEMAQAIREVQSRQEVVRSVLEQVADPDPEGS-SAPRSWKGM	205						
		Dd	172 AVGECDFEVGKERXMKLVKIHPLEK-----VDEATEKKSDGCADSPSSDKEN	222						
		QY	206 NSQVASSILOQ-----PSNLID-----LP-----	222						

```
Db 223 SSOIAQHOKKQETVVKVEDEGRRESINDRARRSPKLPKLSKKGERKWAPPKFLPHKYDVK 282
QY 223 -----PAPELDMWETGPSLT-----FIGHQDIPGVGNHSHGATPPWMIQDEEYIA 267
Db 283 LQNEDKLIISNVPADSLIRTPPNKEIVRYFIRHNAL-----RAGTGENAPWVVEDE--LV 336
QY 268 GNOEIGPSYEEFLKEKEK-----QKLKCLPPDR---VGANFDHSSRTSAGWLPSPGFP 316
Db 337 KKYSLPSKFSDFLLDPYKMYTLNPSTKRKNTGSPDKPKSKKSTDNSSLSS-----PLNP 391
QY 317 RLEKXW-----TPLAQVTS-----TPKUKLQO-----XRSSHIQKXANHAL 351
Db 392 KL--WCHVHLKKSLSGSPKLVKNSKNSKSPHEHLEEMKQMSPNKLIHTNFHIPKKG----445
QY 352 YQLPXCKXP--KSTKPL 366
Db 446 ---PPAKPKGKSDKPL 459

RESULT 12
US-10-703-032-111745
; Sequence 111745, Application US/10703032
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Andersen, Scott E.
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Conner, Timothy W.
; APPLICANT: Cao, Yongwei
; APPLICANT: Masucci, James D.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(53374)B
; CURRENT APPLICATION NUMBER: US/10/703,032
; PRIOR FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 10/020,338
; PRIOR FILING DATE: 2001-12-12
; NUMBER OF SEQ ID NOS: 211164
; SEQ ID NO 111745
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_TA_6163.pep
US-10-703-032-111745

Query Match 4.9%; Score 106.5; DB 6; Length 605;
Best Local Similarity 21.5%; Pred. No. 8.9;
Matches 85; Conservative 53; Mismatches 132; Indels 125; Gaps 20;

QY 20 SIRAHEGPPFCGRGT--FTAASTQRL-----KEAFERLLPOVEAARK--AIRAQVERY 70
Db 93 SSRVNDGICDCGSEEGYNSCTCKNTCWEAGKAAREKLVKKVATYKSGVIRNOVEK- 151
QY 71 VPEHERCCWLCGCEVEHLSHGNLTVLYG-----GLLEHASPEHKKATNKPWNK 124
Db 152 -----AELAIAKDEAELTKLGKEKILQGLVDKL--KEQKRLIEKAESEER 195
QY 125 AEVQMEKEFLVTPQDYARFKSMVKGK-----PSYEKEKDVIKEMAAQIREVQ 174
Db 196 LIKEEERKRIKEVENKQRLKRTIADPAQEAQSDQNTNKAQDESKVVEHHDGDTIDHN 255
QY 175 SRQEVRSV-----LEPOAVPDPPEGSSAPRS-----WKG 204
Db 256 HSPGETSVVEVDIKAGTDDEPPVETSAVPTTEQDPTSVNSEGLSKAELGELVASRTWG 315
QY 205 MN-SQVASSLQO--PSNLDLP--PAPELDMWETGPSLTFIGHDIP-----GVGNHSGAT 255
Db 316 ENVDVVGSKDKKGHEDELDPPEAE-----AFEDHDIPEPVEENYAGVHS---362
QY 256 PPWMIQDEEYIAGNOEI--GPSVEEFLKEKEKQKLK-KLPDPVRVGANFDHSSRTSAG---W 310
Db 363 ---ETEDDRHKFEDELSNVSDOEYVDHDEPDSYKSDDDRKGD--DHSOLTASGQASW 417
```

```
QY 311 LPSFGPRLEXWTPLAVQTSPTPKLKLQXRSSHIQK 345
Db 418 LD-----KIQQTIVQNVLQK 431
```

```
RESULT 13
US-10-450-763-36209
; Sequence 36209, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 36209
; LENGTH: 2246
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (434)..(449)
; OTHER INFORMATION: POLLEN ALLERGEN POA PI SIGNATURE domain identified by
; OTHER INFORMATION: EMATRIX, accession number PR00833H, p-value=3.077e-09, raw score
; OTHER INFORMATION: 2.30
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (4)..(95)
; OTHER INFORMATION: Gag P30 core shell protein domain identified by PFam,
; OTHER INFORMATION: accession name Gag_P30, E-value=1.4e-27, PFam score of 100.2
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2246)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-36209
```

```
Query Match 4.8%; Score 105.5; DB 6; Length 2246;
Best Local Similarity 22.1%; Pred. No. 56;
Matches 98; Conservative 64; Mismatches 135; Indels 147; Gaps 23;

QY 8 RSRIVDPRAAXNS----IRAHEGPPFCGRGTFTAASTQRLKEAFERLLPOVEAARKAIR 63
Db 1419 RELVKEPARAGDRQPEWLEEQGRQFCS---TQAALQAMEREA-EQWGNELERLRAALM 1473
QY 64 AAQVERVYPEHERCCWLCGCEVEHLSHGNLTVLYGGLLEHASPEHKKATNKPWNEN 123
Db 1474 ESQGOQ---QBER-----QQBERE---VARLTQERGRAQADLAL---EKAA-----1510
QY 124 KAEVQMEKEFLVTPQ--DYARFKSMVKGKLSYEEKEDKVIKEM---AAQIREVEQSRQ 178
Db 1511 RAELEMLQNALNQEVFEFATLQELAHALHTEKSGDQDELAKLKGLEAQAQKELEELQOT 1570
QY 179 VVR-----SVLEPQAVPDPPEGSSA 198
Db 1571 VKQLKEQLAKKEKHAHSSGSAQSEAAGRTPTGPKLEALRAEVSKLEQQCQKQQAQDSL 1630
QY 199 PRSKWGNMSQVASSLQPSNLD-LPPAPELDMWETGPSLTFIGH-----QDIP 245
Db 1631 ERS---LEAEASRAERDSDALETTLQGLEEKAQELGHSQSALASQAQRELAARFTKVQD--1685
QY 246 GVGNIHSGATPPWMIQDEEYIAGNOE-----IGPSVEEFL-----KEKEKQKLK 290
Db 1686 -----HSAEDEVKAQ---VARGQEAERKNSLSSLEEEVSIINRQVLEKEGESKEKLR 1737
QY 291 LPPDRVGANFDHSSRTSAGWLPSPGPRLEXWTPLAVQTSPTPKLKLQXRSSHIQKXANHA 350
```

Db 1738 L-----VMAESKSQKLE-----ERLLQAEATASARAAE-RSSALRELEHT 1781  
Qy 351 YQLPXPXGKKPKTKPLL--YLPPK 372  
Db 1782 -----STQALVSELLPAK 1794

RESULT 14  
US-11-097-143-35838  
; Sequence 35838, Application US/11097143  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; TITLE OF INVENTION: DROSOPHILA GENES.  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/11/097,143  
; CURRENT FILING DATE: 2005-04-04  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35838  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: DROSOPHILA

US-11-097-143-35838  
Query Match 4.8%; Score 105; DB 7; Length 669;  
Best Local Similarity 25.1%; Pred. No. 13;  
Matches 65; Conservative 37; Mismatches 99; Indels 58; Gaps 11;  
Qy 122 ENKAEVQMKKFLVTPDYARFKSMVKGDSYEKEDKVIKEMAAQIR-EVEQSRQEV- 179  
Db 407 EDLEMEIK-----HERKELKSHGASTDSMEEDNENETMTVABEHHSDGEVNGEVP 459  
Qy 180 -----VRSVLEPQAVDPPEGSSAPRSWKGNMSQVASSLQOPS---NLDLPP-APELDWM 230  
Db 460 IEERPTTSSEKPSASELPEGNSAPPALKKDVKELOAARQAVSHAVNLLAPPKATEAE-- 517  
Qy 231 ETGFSLFTIGHQDIPGVGNIHSGATPPWMIODEEVIAGNQEIGPSYEEFLKEKEKQKX 290  
Db 518 -----PRTLRSKLSQPP--VVDKK-----SAKQKXKKKKQKQPE 551  
Qy 291 LPPDRVG--ANFDHSSRTSAGWL-----PSFGPRLEXWTP---LAVQTSTPKLKLQXX 338  
Db 552 ASPLKSSDEENHGHRIRTNAGYVTVDEPTKVILIKTSSGMVRVEPTCKQKYFPR 611  
Qy 339 RSSHIQKANHLYQLPXG 357  
Db 612 AAATRKCMGSGARQDPPG 630

RESULT 15  
US-10-940-774A-6676  
; Sequence 6676, Application US/10940774A  
; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/10/940,774A  
; CURRENT FILING DATE: 2004-09-15  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6676  
; LENGTH: 753  
; TYPE: PRT  
; ORGANISM: Human  
; US-10-940-774A-6676

Query Match 4.7%; Score 103.5; DB 6; Length 753;  
Best Local Similarity 22.0%; Pred. No. 20;  
Matches 84; Conservative 48; Mismatches 149; Indels 101; Gaps 19;  
Qy 19 NSIRAH-----EGPFPCG-RGTTAASTQRLK-----EAFERLLPQVE 56  
Db 185 HSLRIHVHVHTKGPEDCVQCEKAFNTLYRLKAHQRLHTGKTFCNCESEGCSEKVFYFTLS 244  
Qy 57 AAKATRAAOVER-----YVPEH--ERCWCCLCCGE----- 86  
Db 245 DLRRKHINTHTGKPFRCDDHCGCKAFAAASHHLKTHVTHTGERTPFPCPSNGCEKTFSTQY 304  
Qy 87 -VREHL-SHGNLTLYLGLLEHLASPEHKK-----ATNKFWMENKAEVQMKKFL 134  
Db 305 SLKSHMKGNKSHYNALPQNGSEDTNHSLSLSDLSLSTSELRENSSTTQGDLS 364  
Qy 135 VTPQDYARFKSMVKGDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVDPPEE 194  
Db 365 ISP---AIIFESM-----FQNSDDTAIQEDPQQTASLTESFNGDAESVSD---VP-PST 411  
Qy 195 GSSAPRS-----WKGMSQVASSLQOPSNLDLPPAPELDWMETGPSLTFIHHQDIPGVGN 249  
Db 412 GNSASLSLPLVLQPLGLSEPPQPLLPASAPSAPPAPSL-----GP-----GSQQ-AAFGN 460  
Qy 250 IHSGATPPWMIODEEY-IAGNQEIGPSYEEFLKEKEKQKLLKLPDRVGFANFDHSSRTSA 308  
Db 461 -----PPALLQPEVPVPHSTOFAANHQBFLPHQAPQ-PIVPGLSVVAGASASAAVA 513  
Qy 309 GWLPFGPRLEXWTP-LAVQTST 330  
Db 514 SAVAAPAPPOSTTEPLPAMVQT 535

Search completed: June 7, 2005, 12:21:18  
Job time : 22.7652 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:52:21 ; Search time 186.04 Seconds  
(without alignments)  
2618.037 Million cell updates/sec

Title: US-09-155-676B-5

Perfect score: 2190

Sequence: 1 IGTVRRSRRIVDPRAXNS.....RGLESSLSIFSTSGRGTRHF 417

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6959266 seqs, 1168006243 residues

Total number of hits satisfying chosen parameters: 6959266

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending Patents AA.Main.\*
- 1: /cgn2\_6/ptodata/1/paa/PTCTUS\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*
  - 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*
  - 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*
  - 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*
  - 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*
  - 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*
  - 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*
  - 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*
  - 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*
  - 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*
  - 21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*
  - 22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep.\*
  - 23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*
  - 24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep.\*
  - 25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*
  - 26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*
  - 27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*
  - 28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*
  - 29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*
  - 30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*
  - 31: /cgn2\_6/ptodata/1/paa/US105\_COMB.pep.\*
  - 32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*
  - 33: /cgn2\_6/ptodata/1/paa/US107\_COMB.pep.\*
  - 34: /cgn2\_6/ptodata/1/paa/US108\_COMB.pep.\*
  - 35: /cgn2\_6/ptodata/1/paa/US109\_COMB.pep.\*
  - 36: /cgn2\_6/ptodata/1/paa/US110\_COMB.pep.\*
  - 37: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2178	99.5	417	15	US-09-155-676-5 Sequence 5, Appli
2	2178	99.5	417	15	US-09-155-676A-5 Sequence 5, Appli
3	972.5	44.4	211	27	US-10-170-2058-36800 Sequence 9118, Ap
4	972.5	44.4	212	37	US-60-453-050-9118 Sequence 9118, Ap
5	972.5	44.4	212	37	US-60-453-135-9118 Sequence 9118, Ap
6	972.5	44.4	212	37	US-60-466-412-9118 Sequence 9118, Ap
7	795	36.3	1115	24	US-09-915-543-17 Sequence 17, Appl
8	795	36.3	1115	29	US-10-322-579-17 Sequence 17, Appl
9	795	36.3	1115	22	US-10-564-859-17 Sequence 17, Appl
10	731.5	33.4	216	22	US-09-758-463-1029 Sequence 1029, Ap
11	731.5	33.4	216	28	US-10-217-607-1029 Sequence 1029, Ap
12	670	30.6	156	1	PCT-US01-08631-49939 Sequence 49939, A
13	554.5	25.3	741	1	PCT-US01-08631-49940 Sequence 49940, A
14	517	23.6	189	1	PCT-US04-12047-1081 Sequence 1081, Ap
15	503	23.0	312	1	PCT-US01-08631-44039 Sequence 44039, A
16	309.5	14.1	65	1	PCT-US01-01354-17236 Sequence 17236, A
17	309.5	14.1	65	22	US-09-764-905-17236 Sequence 17236, A
18	309.5	14.1	65	26	US-10-092-399-17236 Sequence 17236, A
19	198	9.0	104	1	PCT-US98-23829-2 Sequence 2, Appli
20	198	9.0	104	15	US-09-185-936-2 Sequence 2, Appli
21	198	9.0	104	21	US-09-746-783-206 Sequence 206, App
22	198	9.0	104	32	US-10-689-742-206 Sequence 206, App
23	183	8.4	71	17	US-09-330-337-292 Sequence 292, App
24	148	6.8	389	30	US-10-424-599-209186 Sequence 209186, A
25	134	6.1	423	21	US-09-708-427-26860 Sequence 26860, A
26	117.5	5.4	443	37	US-60-556-841-5931 Sequence 5931, Ap
27	117.5	5.4	1040	30	US-10-425-115-308861 Sequence 308861, A
28	116	5.3	367	28	US-10-219-999-39253 Sequence 39253, A
29	116	5.3	367	30	US-10-425-114-41619 Sequence 41619, A
30	116	5.3	367	30	US-10-425-114A-41619 Sequence 41619, A
31	116	5.3	367	37	US-60-312-544-8751 Sequence 8751, Ap
32	116	5.3	368	37	US-60-324-109-26189 Sequence 26189, A
33	116	5.3	369	28	US-10-219-999-59765 Sequence 59765, A
34	116	5.3	369	30	US-10-425-114-69227 Sequence 69227, A
35	116	5.3	369	30	US-10-425-114A-69227 Sequence 69227, A
36	116	5.3	653	28	US-10-219-999-40981 Sequence 40981, A
37	116	5.3	653	30	US-10-425-114-51821 Sequence 51821, A
38	116	5.3	653	30	US-10-425-114A-51821 Sequence 51821, A
39	116	5.3	653	37	US-60-324-109-17151 Sequence 17151, A
40	112	5.1	124	28	US-10-219-999-49996 Sequence 49996, A
41	112	5.1	124	28	US-10-219-999-50911 Sequence 50911, A
42	112	5.1	160	30	US-10-425-115-219898 Sequence 219898, A
43	112	5.1	253	30	US-10-425-114-62537 Sequence 62537, A
44	112	5.1	253	30	US-10-425-114-67550 Sequence 67550, A
45	112	5.1	253	30	US-10-425-114A-62537 Sequence 62537, A

ALIGNMENTS

RESULT 1  
US-09-155-676-5  
Sequence 5, Application US/09155676  
GENERAL INFORMATION:  
APPLICANT: WALLACH, David  
APPLICANT: MALININ, Nikolai  
APPLICANT: BOLDIN, Mark  
APPLICANT: KOVALENKO, Andrei  
APPLICANT: METT, Igor  
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
STREET: 624 Ninth Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20001

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-155-676-5
```

```
Query Match 99.5%; Score 2178; DB 15; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.1e-186;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGVTWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
DB 1 IGVTWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
QY 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKF 120
DB 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKF 120
QY 121 WENKAEVQMKERFLVTPQDYARFKSMVKGKLDSEYEEKDKVTKENAAQIREVEQSRQEV 180
DB 121 WENKAEVQMKERFLVTPQDYARFKSMVKGKLDSEYEEKDKVTKENAAQIREVEQSRQEV 180
QY 181 RSVLEPQAVDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLITFTG 240
DB 181 RSVLEPQAVDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLITFTG 240
QY 241 HDIFGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKQKLLKLPDRVGANF 300
DB 241 HDIFGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKQKLLKLPDRVGANF 300
QY 301 DHSSTTSAGWLPSPGRLEXTWPLAVQSTPKLKLQQRSSHIOKKANHALYQLPXGKXP 360
DB 301 DHSSTTSAGWLPSPGRLEXTWPLAVQSTPKLKLQQRSSHIOKKANHALYQLPXGKXP 360
QY 361 KSTKPLLYLPKPFVILVRKQTYFIFXENKVLFFGLRGLESSLSIPSTSRGGRTHF 417
DB 361 KSTKPLLYLPKPFVILVRKQTYFIFXENKVLFFGLRGLESSLSIPSTSRGGRTHF 417
```

## RESULT 2

```
US-09-155-676A-5
; Sequence 5, Application US/09155676A
; GENERAL INFORMATION:
; APPLICANT: WALLACH, David
; APPLICANT: MALININ, Nikolai
```

```
APPLICANT: BOLDIN, Mark
APPLICANT: KOVALENKO, Andrei
APPLICANT: METT, Igor
TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED
TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/155,676A
FILING DATE: 04-JAN-1999
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IL97/00117
FILING DATE: 01-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117800
FILING DATE: 02-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 119133
FILING DATE: 26-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-155-676A-5

Query Match 99.5%; Score 2178; DB 15; Length 417;
Best Local Similarity 100.0%; Pred. No. 2.1e-186;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IGVTWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
DB 1 IGVTWRSRIVDPRAAXNSIRAHGPPFCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
QY 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKF 120
DB 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKF 120
QY 121 WENKAEVQMKERFLVTPQDYARFKSMVKGKLDSEYEEKDKVTKENAAQIREVEQSRQEV 180
DB 121 WENKAEVQMKERFLVTPQDYARFKSMVKGKLDSEYEEKDKVTKENAAQIREVEQSRQEV 180
QY 181 RSVLEPQAVDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLITFTG 240
DB 181 RSVLEPQAVDPPEGSSAPRSWKGMNSQVASSLQOPSNLDLPPAPELDMETGPSLITFTG 240
QY 241 HDIFGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKQKLLKLPDRVGANF 300
DB 241 HDIFGVGNHSGATPPWMIQDEEYIAGNQEIGPSYEEFLKEKEKQKLLKLPDRVGANF 300
QY 301 DHSSTTSAGWLPSPGRLEXTWPLAVQSTPKLKLQQRSSHIOKKANHALYQLPXGKXP 360
```

Db 301 DHSRTSAGWLPSPGRLEKXTPTLAVQTSTPKLKLQQRSSHIQKANHLYQLPXGKXP 360  
Qy 361 KSTRPLLYLPKFFIIVFLRKQTYSFIXFNKVLFFGLRGLSSLSIPSTSRGGRTHF 417  
Db 361 KSTRPLLYLPKFFIIVFLRKQTYSFIXFNKVLFFGLRGLSSLSIPSTSRGGRTHF 417

RESULT 3  
US-10-170-205E-36800  
; Sequence 36800, Application US/10170205E  
; GENERAL INFORMATION:  
; APPLICANT: ADAMS, Mark  
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN  
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF  
; FILE REFERENCE: CL001381  
; CURRENT APPLICATION NUMBER: US/10/170,205E  
; CURRENT FILING DATE: 2002-06-13  
; NUMBER OF SEQ ID NOS: 40312  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 36800  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 27; Length 211;  
Best Local Similarity 74.5%; Pred. No. 2.5e-78;  
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87  
Db 14 FFCGRGHVYSRKQORQLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147  
Db 74 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207  
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 171

Qy 208 QVASSLQOPSNLDPAPPELDMETGSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNOEIGPSYEEFLKE 282  
Db 197 GNOEIGPSYEEFLKE 211

US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 27; Length 211;  
Best Local Similarity 74.5%; Pred. No. 2.5e-78;  
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87  
Db 14 FFCGRGHVYSRKQORQLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147  
Db 74 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207  
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 171

Qy 208 QVASSLQOPSNLDPAPPELDMETGSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNOEIGPSYEEFLKE 282  
Db 197 GNOEIGPSYEEFLKE 211

US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 27; Length 211;  
Best Local Similarity 74.5%; Pred. No. 2.5e-78;  
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87  
Db 14 FFCGRGHVYSRKQORQLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147  
Db 74 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207  
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 171

Qy 208 QVASSLQOPSNLDPAPPELDMETGSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNOEIGPSYEEFLKE 282  
Db 197 GNOEIGPSYEEFLKE 211

US-10-170-205E-36800

Query Match 44.4%; Score 972.5; DB 37; Length 212;  
Best Local Similarity 74.5%; Pred. No. 2.5e-78;  
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87  
Db 14 FFCGRGHVYSRKQORQLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147  
Db 74 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207  
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 171

Qy 208 QVASSLQOPSNLDPAPPELDMETGSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNOEIGPSYEEFLKE 282  
Db 197 GNOEIGPSYEEFLKE 211

US-60-453-135-9118

RESULT 5  
US-60-453-135-9118  
; Sequence 9118, Application US/60453135  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: IAKOUBOVA, Olga  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001456  
; CURRENT APPLICATION NUMBER: US/60/453,135  
; CURRENT FILING DATE: 2003-03-10  
; NUMBER OF SEQ ID NOS: 82762  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9118  
; LENGTH: 212  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(212)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-60-453-135-9118

Query Match 44.4%; Score 972.5; DB 37; Length 212;  
Best Local Similarity 74.5%; Pred. No. 2.5e-78;  
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

Qy 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87  
Db 14 FFCGRGHVYSRKQORQLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 73

Qy 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 147  
Db 74 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFKLVTPDYARFKKSM 133

Qy 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207  
Db 134 VKGLDSYEEDKVKIKEMAAQIREVQSRQVRSVLEPQAVDPPEEGSSAPRSWKGMNS 171

Qy 208 QVASSLQOPSNLDPAPPELDMETGSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 172 -----DIPGVGNIHSGATPPPMIODEEYIA 196

Qy 268 GNOEIGPSYEEFLKE 282  
Db 197 GNOEIGPSYEEFLKE 211

US-60-453-135-9118

```
RESULT 6
US-60-466-412-9118
; Sequence 9118, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: TAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9118
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(212)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-466-412-9118

Query Match 44.4%; Score 972.5; DB 37; Length 212;
Best Local Similarity 74.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 1; Mismatches 7; Indels 57; Gaps 1;

QY 28 FFCGRGTFTAASTQROLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 87
DB 14 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 73

QY 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 147
DB 74 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 133

QY 148 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207
DB 134 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLE-----171

QY 208 QVASSLQOPSNDLPPAPELDWMETGPSLTFIGHQDIPGVGNIHSGATPPMMIQDEEYIA 267
DB 172 -----DIPGVGNIHSGATPPMMIQDEEYIA 196

QY 268 GNOEIGPSYEEFLKE 282
DB 197 GNOEIGPSYEEFLKE 211

RESULT 7
US-09-915-543-17
; Sequence 17, Application US/09915543
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/09/915,543
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-09-915-543-17

Query Match 36.3%; Score 795; DB 24; Length 1115;
Best Local Similarity 55.6%; Pred. No. 2.9e-61;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGTFTAASTQROLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 87
DB 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 887

QY 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 147
DB 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 947

QY 148 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207
DB 948 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLE-----TGPPR-----990

QY 208 QVASSLQOPSNDLPPAPELDWMETGPSLTFIGHQDIPGVGNIHSGATPPMMIQDEEYIA 267
DB 991 -YALTVRSPAVLSR-----RTLSKGAFPPQ--TPEAHPQ 1021

Query Match 36.3%; Score 795; DB 29; Length 1115;
Best Local Similarity 55.6%; Pred. No. 2.9e-61;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGTFTAASTQROLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 87
DB 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 887

QY 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 147
DB 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 947

QY 148 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207
DB 948 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLE-----TGPPR-----990

QY 208 QVASSLQOPSNDLPPAPELDWMETGPSLTFIGHQDIPGVGNIHSGATPPMMIQDEEYIA 267
DB 991 -YALTVRSPAVLSR-----RTLSKGAFPPQ--TPEAHPQ 1021
```

```
Best Local Similarity 55.6%; Pred. No. 2.9e-61;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGTFTAASTQROLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 87
DB 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 887

QY 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 147
DB 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 947

QY 148 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207
DB 948 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLE-----TGPPR-----990

QY 208 QVASSLQOPSNDLPPAPELDWMETGPSLTFIGHQDIPGVGNIHSGATPPMMIQDEEYIA 267
DB 991 -YALTVRSPAVLSR-----RTLSKGAFPPQ--TPEAHPQ 1021

QY 268 GNOEIGPSYEEFLKEKQKLK-KLPDDRVGANFDHSSRTSAGWLPSPGRLXWTPLAV 326
DB 1022 ARCLCAPRRGALKPEPPGRTTLKLVPHHTTKARPHAAKTSP-----RPRCTROAP---1072

QY 327 QTSTPKLKL---QQXRSSHIOKKA 347
DB 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 8
US-10-322-579-17
; Sequence 17, Application US/10322579
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; FILE REFERENCE: Q60361
; CURRENT APPLICATION NUMBER: US/10/322,579
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-10-322-579-17

Query Match 36.3%; Score 795; DB 29; Length 1115;
Best Local Similarity 55.6%; Pred. No. 2.9e-61;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGTFTAASTQROLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 87
DB 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCGCEV 887

QY 88 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 147
DB 888 REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKFVLTPQDYARFKKSM 947

QY 148 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLEPQAVDPPEEGSSAPRSWKGMNS 207
DB 948 VKGLDSYEKEDKVIKEMAAQIREVQSRQEVVRSVLE-----TGPPR-----990

QY 208 QVASSLQOPSNDLPPAPELDWMETGPSLTFIGHQDIPGVGNIHSGATPPMMIQDEEYIA 267
DB 991 -YALTVRSPAVLSR-----RTLSKGAFPPQ--TPEAHPQ 1021
```



```
QY 268 GNOEIGSYERFLKEKEKQK-LKLPDRVGANFDHSRTSAGWLPFGPRLXKWTPLAV 326
DB 1022 ARCLCAPRGALKPEPPGRTLKLGVPHTTKRPHAAKTSF-----RPRCTRQAP--- 1072
QY 327 QTSTPKLKL---QXRSRSHIOKKA 347
DB 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 9
US-10-664-859-17
; Sequence 17, Application US/10664859
; GENERAL INFORMATION:
; APPLICANT: BASLER, Konrad
; APPLICANT: BRUNNER, Erich
; APPLICANT: FROESCH, Barbara
; APPLICANT: KRAMPS, Thomas
; APPLICANT: PETER, Oliver
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY
; FILE REFERENCE: Q60361
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON
; CURRENT APPLICATION NUMBER: US/10/664,859
; PRIOR FILING DATE: 2003-09-22
; PRIOR APPLICATION NUMBER: US/09/915,543
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/221,502
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1115
; TYPE: PRT
; ORGANISM: Human lgs-1
US-10-664-859-17

Query Match 36.3%; Score 795; DB 32; Length 1115;
Best Local Similarity 55.6%; Pred. No. 2.9e-61;
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;

QY 28 FFCGRGFTTAASRQRLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
DB 828 FFCGRGHVYSRKHQRQKQKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 887
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 147
DB 888 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 947
QY 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMNS 207
DB 948 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLE-----TGPPR----- 990
QY 208 QVASSLOQPSNLDLPAPPELDMETGSLTFIGHQDIPGVGNHSGATPPMIODESYIA 267
DB 991 -YALTVRSPAVLSR-----RTLKSGAPFPQ--TPEAHPQ 1021
QY 268 GNOEIGSYERFLKEKEKQK-LKLPDRVGANFDHSRTSAGWLPFGPRLXKWTPLAV 326
DB 1022 ARCLCAPRGALKPEPPGRTLKLGVPHTTKRPHAAKTSF-----RPRCTRQAP--- 1072
QY 327 QTSTPKLKL---QXRSRSHIOKKA 347
DB 1073 -NKTQSLQLAGKARKTALHLQTKA 1095

RESULT 10
US-09-758-463-1029
; Sequence 1029, Application US/09758463
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM011
; CURRENT APPLICATION NUMBER: US/09/758,463
```

```
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1029
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-463-1029

Query Match 33.4%; Score 731.5; DB 22; Length 216;
Best Local Similarity 78.8%; Pred. No. 1.3e-56;
Matches 141; Conservative 3; Mismatches 12; Indels 23; Gaps 1;

QY 28 FFCGRGFTTAASRQRLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
DB 61 FFCGRGHVYSRKHQRQKQKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 120
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 147
DB 121 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 180
QY 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMN 206
DB 181 VKGLDSYEEDKVKIKEMAAQIR-----EWSRADRRWFGLS 216

RESULT 11
US-10-217-607-1029
; Sequence 1029, Application US/10217607
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM011CIN
; CURRENT APPLICATION NUMBER: US/10/217,607
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 09/758,463
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1029
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-607-1029

Query Match 33.4%; Score 731.5; DB 28; Length 216;
Best Local Similarity 78.8%; Pred. No. 1.3e-56;
Matches 141; Conservative 3; Mismatches 12; Indels 23; Gaps 1;

QY 28 FFCGRGFTTAASRQRLKEAFERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 87
DB 61 FFCGRGHVYSRKHQRQKQKEALERLLPQVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEV 120
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 147
DB 121 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWENKAQVOMKEKFLVTPDYARFKKSM 180
QY 148 VKGLDSYEEDKVKIKEMAAQIREVQSRQEVRSVLEPQAVDPPEGSSAPRSWKGMN 206
DB 181 VKGLDSYEEDKVKIKEMAAQIR-----EWSRADRRWFGLS 216

RESULT 12
PCT-US01-08631-49939
; Sequence 49939, Application PC/TUS0108631
```

```

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 49939
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)-(156)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-49939

Query Match 30.6%; Score 670; DB 1; Length 156;
Best Local Similarity 77.9%; Pred. No. 2.7e-51;
Matches 134; Conservative 4; Mismatches 12; Indels 22; Gaps 4;

QY 6 WRSRIVDPAAXNSIRAHEGPFPCGRTFTAASTORQLKEAFERLLLPQVEAARKAIRAA 65
DB 5 WSRARL-QQAPA-----AAGGFG--GEAPAPGA-----EVEAARKAIRAA 42
QY 66 QERYVPERHCWCCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKA 125
DB 43 QERYVPERHCWCCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKA 102
QY 126 EVQMKKEFLVTPQDYARFKKSMVKGGLDSYEEKEDKVIKEMAAQIREVEQSQR 177
DB 103 EVQMKKEFLVTPQDYARFKKSMVKGGLDSYEEKEDKVIKEMAAQIREVEQSQR 154

RESULT 13
PCT-US01-08631-49940
; Sequence 49940, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 49940
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (632)...(650)
; OTHER INFORMATION: BURKITT'S LYMPHOMA RECEPTOR SIGNATURE domain identified by
; OTHER INFORMATION: EMATRIX, accession number PR00564E, p-value=1.918e-23, raw score
; OTHER INFORMATION: 12.84
; NAME/KEY: DOMAIN
; LOCATION: (502)...(736)
; OTHER INFORMATION: 7 transmembrane receptor domain identified by Pfam, accession
; OTHER INFORMATION: name 7cm_1, E-value=5.2e-81, Pfam score of 259.5
PCT-US01-08631-49940

Query Match 25.3%; Score 554.5; DB 1; Length 741;
Best Local Similarity 40.7%; Pred. No. 7e-40;
Matches 135; Conservative 18; Mismatches 52; Indels 127; Gaps 11;

```

```

QY 66 QERYVPERHCWCCLCCGCEVREHLSHGNTLVLYGGLLEHL-----ASPEHKK----- 114
DB 175 KERYVPERHCWCCLCCGCEVREHLSHGNTLVLYGGLLEHLARTASLVAAPNWKQFRCL 234
QY 115 ATNKFWMENKAEVQMKKEFLVTPQDYARFKKSMVKGGLDSYEEKEDKVIKEMAAQIREVEQ 174
DB 235 STGK--WINNG-----IVFQEDRI-----IDTYDTDES-----OK 263
QY 175 SRQEVVRSVLEPQAVPDPEEGSSAPRSKGMNSQVASSLQOPSNLDLPPAPELDWMTGP 234
DB 264 SR----- 265
QY 235 SLTFIGHQDIPGVGNIHSGATPPMWIQDEEYIAGNOBIPGSYEEFLKEKEKQKLKLPDP 294
DB 266 -----PG-----RRGATPPMWIQDEEYIAGNOBIPGSYEEFLKEKEKQKLKLPDP 311
QY 295 RVGANFDHSSRTSAGWLPSPFGFLXWTPLAVOTSTPKLKLQQRSSHIQKANHALYQL 354
DB 312 RVGANFDHSSRTSAGWLPSPFG-----RVWN-----NGRRWQSRAPERGEVEVPWSEQT 359
QY 355 PXGKPKSTK-PLLYLPKPFILVFLRKQTVS 385
DB 360 PDPHTSRSLRIPAEVAPRSKRRCVLERKQPYS 391

RESULT 14
PCT-US04-12047-1081
; Sequence 1081, Application PC/TUS0412047
; GENERAL INFORMATION:
; APPLICANT: FIVE PRIME THERAPEUTICS, INC.
; TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF
; TITLE OF INVENTION: THEIR USE
; FILE REFERENCE: 08940.0021-00304
; CURRENT APPLICATION NUMBER: PCT/US04/12047
; CURRENT FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 1464
; PRIOR APPLICATION NUMBER: US 60/463,732
; PRIOR FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/463,708
; PRIOR FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: US 60/467,230
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/467,199
; PRIOR FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: US 60/493,573
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US 60/493,577
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US 60/486,480
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/486,446
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US 60/471,306
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/471,336
; PRIOR FILING DATE: 2003-05-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1081
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-12047-1081

Query Match 23.6%; Score 517; DB 1; Length 189;
Best Local Similarity 64.1%; Pred. No. 2.1e-37;
Matches 107; Conservative 18; Mismatches 36; Indels 6; Gaps 3;

QY 133 FLVTPQDYARFKKSMVKGGLDSYEEKEDKVIKEMAAQIREVEQSQRVRSVLE---PQAV 189
DB 24 YLASP-DYARFKKSLVEDLGSYEEKEDKVIKEMAAQICEVEQSWQVQSVLEVGFFRI 82

```

QY 190 PDPEGSSAPRSWKGMNSQVASSLQOPSNLDLPP--APELDWMTGPSLTETIGHQDIPGV 247  
Db 83 QTPSASAGHFSWKLKDEQPGSFQPTAGLQGPATAPELDWMTGLSPIFIGHQDIPGV 142  
QY 248 GNIHSGATPPWMIQDEEYIAGNQETGSPSYEFLEKEKOKLKLPPD 294  
Db 143 GNIHSGATPPWMIQDEECISGNQQTGSPSYKEFIKEKEKELKELSPN 189

RESULT 15

PCT-US01-08631-44039  
; Sequence 44039, Application PC/TUS0108631  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-049  
; CURRENT APPLICATION NUMBER: PCT/US01/08631  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 44039  
; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-08631-44039

Query Match 23.0%; Score 503; DB 1; Length 312;  
Best Local Similarity 49.3%; Pred. No. 8.2e-36;  
Matches 113; Conservative 14; Mismatches 38; Indels 64; Gaps 6;  
QY 40 TOROLKEAFERL-----LPOVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEVREHL 91  
Db 108 TOPRLPAADWMPAATAITGTDVEAARKAIRAAQVERVYVPEHERCCWCLCCGCEVREHL 167  
QY 92 SHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKKFLVTPQDYARF----- 143  
Db 168 SHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKKFLVTPQDYARSHLYGTPMG 227  
QY 144 ----KKSVMKGL-----DSYEKEDKVIKEMAAQIREVEQSRQEVVRSV-LEPOA 188  
Db 228 SAVFSQACLEALLIIVPMGACGVSQEEES-----PARGSKDEPGEQVELKEEA 276  
QY 189 VPDPEGSSAPRSWKGMNSQVASSIQQPSNLDLPPAPELDWMTGPSLT 237  
Db 277 EAPVEDGSQPP-----PPEPKGDATPEGEKAT 303

Search completed: June 7, 2005, 12:19:34  
Job time : 188.04 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 12:02:31 ; Search time 52.7607 Seconds  
(without alignments)  
2846.907 Million cell updates/sec

Title: US-09-155-676B-5  
Perfect score: 2190  
Sequence: 1 IGTVRRRSRIVDPAAXNS.....RGLESSISPTSRGGRTHF 417

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1599520 seqs, 360203123 residues

Total number of hits satisfying chosen parameters: 1599520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/FCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	795	36.3	1115	9 US-09-915-543-17	Sequence 17, Appl
2	795	36.3	1115	14 US-10-322-579-17	Sequence 17, Appl
3	795	36.3	1115	15 US-10-664-859-17	Sequence 17, Appl
4	198	9.0	104	10 US-09-746-783-206	Sequence 206, App
5	148	6.8	389	15 US-10-424-599-209186	Sequence 209186,
6	117.5	5.4	1040	16 US-10-425-115-306861	Sequence 306861,
7	116	5.3	367	15 US-10-425-114-41619	Sequence 41619, A
8	116	5.3	369	15 US-10-425-114-69227	Sequence 69227, A
9	116	5.3	653	15 US-10-425-114-51821	Sequence 51821, A
10	112	5.1	160	16 US-10-425-115-219898	Sequence 219898,
11	112	5.1	253	15 US-10-425-114-62537	Sequence 62537, A
12	112	5.1	253	15 US-10-425-114-67550	Sequence 67550, A

13	110.5	5.0	1878	15	US-10-607-631-20	Sequence 20, Appl
14	109.5	5.0	1109	16	US-10-437-963-133129	Sequence 133129,
15	109	5.0	4952	15	US-10-051-874-56	Sequence 56, Appl
16	109	5.0	5008	15	US-10-051-874-166	Sequence 166, App
17	109	5.0	5159	15	US-10-085-198-112	Sequence 112, App
18	109	5.0	5262	15	US-10-051-874-165	Sequence 165, App
19	107	4.9	447	16	US-10-425-115-368903	Sequence 368903,
20	106	4.8	270	14	US-10-105-021-2	Sequence 2, Appli
21	105.5	4.8	1879	15	US-10-607-631-8	Sequence 8, Appli
22	105	4.8	583	16	US-10-425-115-294295	Sequence 294295,
23	105	4.8	620	15	US-10-425-114-62454	Sequence 62454, A
24	105	4.8	2701	14	US-10-171-311-83	Sequence 83, Appl
25	104.5	4.8	526	15	US-10-094-749-1861	Sequence 1861, Ap
26	104.5	4.8	1094	14	US-10-032-585-7554	Sequence 7554, Ap
27	104.5	4.8	2093	16	US-10-437-963-128540	Sequence 128540,
28	104	4.7	1525	9	US-09-839-479-68	Sequence 68, Appl
29	104	4.7	1525	15	US-10-376-537-69	Sequence 69, Appl
30	104	4.7	1525	15	US-10-702-148-68	Sequence 68, Appl
31	104	4.7	5262	15	US-10-051-874-167	Sequence 167, App
32	103.5	4.7	753	16	US-10-755-889-358	Sequence 358, App
33	103.5	4.7	3224	10	US-09-315-355-34	Sequence 34, Appl
34	103	4.7	393	9	US-09-925-397-825	Sequence 825, App
35	103	4.7	905	13	US-10-114-893-127	Sequence 127, App
36	103	4.7	1056	15	US-10-275-595A-30	Sequence 30, Appl
37	103	4.7	1527	9	US-09-839-479-27	Sequence 27, Appl
38	103	4.7	1527	15	US-10-376-537-27	Sequence 27, Appl
39	103	4.7	1527	15	US-10-702-148-27	Sequence 27, Appl
40	103	4.7	1531	9	US-09-839-479-29	Sequence 29, Appl
41	103	4.7	1531	15	US-10-376-537-29	Sequence 29, Appl
42	103	4.7	1531	15	US-10-702-148-29	Sequence 29, Appl
43	102.5	4.7	563	16	US-10-425-115-286927	Sequence 286927,
44	102.5	4.7	964	16	US-10-437-963-181208	Sequence 181208,
45	102.5	4.7	1009	8	US-08-987-689A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-915-543-17  
; Sequence 17, Application US/09915543  
; Publication No. US20020086986A1  
; GENERAL INFORMATION:  
; APPLICANT: BASLER, Konrad  
; APPLICANT: BRUNNER, Erich  
; APPLICANT: FROESCH, Barbara  
; APPLICANT: KRAMPS, Thomas  
; APPLICANT: PETER, Oliver  
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY  
; FILE REFERENCE: Q60361  
; CURRENT APPLICATION NUMBER: US/09/915,543  
; CURRENT FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: 60/221,502  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 1115  
; TYPE: PRT  
; ORGANISM: Human lgs-1  
US-09-915-543-17

Query Match	36.3%	Score	795	DB	9	Length	1115
Best Local Similarity	55.6%	Pred. No.	3.5e-58				
Matches	180	Conservative	17	Mismatches	67	Indels	60
Gaps	8						
Qy	28	FFCGRGTTAASTORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV	87				
Db	828	FFCGRGHVSRKHQRLKEALERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV	887				
Qy	88	REHLSHGNLTVLYGGLLEHLASPEHKATKNFHWENKAQVOMKEKFLVTPDYARFKKSM	147				

Db 888 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKAEVQMKFKLVTPQDYARFKKSM 947  
QY 148 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPPEBGSAPRSWKGMNS 207  
Db 948 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990  
QY 208 QVASSLOQPSNLDLPPAPELDMWETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 991 -YALTVRSPAVLSR-----RTLKSGAPPPQ--TPEAHPQ 1021  
QY 268 GNOETGPSYEBFLKEKEQKLIK-KLPDPVRGANGFHDHSSRTSAGWLPSPGPRLEXWTPPLAV 326  
Db 1022 ARCLCAPRRGALKPEPPGRTKLGVPPTHTRKARPHAAKTSP-----RPRCTROAP--- 1072  
QY 327 QTSTPKLKL---QOXRSSHIOKKA 347  
Db 1073 -NKTOSLQLAGKARKTALHLQTKA 1095  
RESULT 2  
US-10-322-579-17  
; Sequence 17, Application US/10322579  
; Publication No. US20030114413A1  
; GENERAL INFORMATION:  
; APPLICANT: BASLER, Konrad  
; APPLICANT: BRUNNER, Erich  
; APPLICANT: FROESCH, Barbara  
; APPLICANT: KRAMPS, Thomas  
; APPLICANT: PETER, Oliver  
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON  
; FILE REFERENCE: Q60361  
; CURRENT APPLICATION NUMBER: US/10/322,579  
; CURRENT FILING DATE: 2002-12-19  
; PRIOR APPLICATION NUMBER: US/09/915,543  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 17  
; LENGTH: 1115  
; TYPE: PRT  
; ORGANISM: Human lgs-1  
US-10-322-579-17  
Query Match 36.3%; Score 795; DB 14; Length 1115;  
Best Local Similarity 55.6%; Pred. No. 3.5e-58;  
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;  
QY 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV 87  
Db 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV 887  
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKAEVQMKFKLVTPQDYARFKKSM 147  
Db 888 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKAEVQMKFKLVTPQDYARFKKSM 947  
QY 148 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPPEBGSAPRSWKGMNS 207  
Db 948 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990  
QY 208 QVASSLOQPSNLDLPPAPELDMWETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 991 -YALTVRSPAVLSR-----RTLKSGAPPPQ--TPEAHPQ 1021  
QY 268 GNOETGPSYEBFLKEKEQKLIK-KLPDPVRGANGFHDHSSRTSAGWLPSPGPRLEXWTPPLAV 326  
Db 1022 ARCLCAPRRGALKPEPPGRTKLGVPPTHTRKARPHAAKTSP-----RPRCTROAP--- 1072  
QY 327 QTSTPKLKL---QOXRSSHIOKKA 347  
Db 1073 -NKTOSLQLAGKARKTALHLQTKA 1095

RESULT 3  
US-10-664-859-17  
; Sequence 17, Application US/10664859  
; Publication No. US20040038901A1  
; GENERAL INFORMATION:  
; APPLICANT: BASLER, Konrad  
; APPLICANT: BRUNNER, Erich  
; APPLICANT: FROESCH, Barbara  
; APPLICANT: KRAMPS, Thomas  
; APPLICANT: PETER, Oliver  
; TITLE OF INVENTION: ESSENTIAL DOWNSTREAM COMPONENT OF THE WINGLESS SIGNALING PATHWAY  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC APPLICATIONS BASED THEREON  
; FILE REFERENCE: Q60361  
; CURRENT APPLICATION NUMBER: US/10/664,859  
; CURRENT FILING DATE: 2003-09-22  
; PRIOR APPLICATION NUMBER: US/09/915,543  
; PRIOR FILING DATE: 2001-07-27  
; PRIOR APPLICATION NUMBER: 60/221,502  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 17  
; LENGTH: 1115  
; TYPE: PRT  
; ORGANISM: Human lgs-1  
US-10-664-859-17  
Query Match 36.3%; Score 795; DB 15; Length 1115;  
Best Local Similarity 55.6%; Pred. No. 3.5e-58;  
Matches 180; Conservative 17; Mismatches 67; Indels 60; Gaps 8;  
QY 28 FFCGRGTFTAATORQLKEAFERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV 87  
Db 828 FFCGRGHVYSRKHQRLKEALERLLPQVEAARKAIRAAQVRYVPEHERCCWCLCCGCEV 887  
QY 88 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKAEVQMKFKLVTPQDYARFKKSM 147  
Db 888 REHLSHGNTLVLYGGLLEHLASPEHKATNKFWWENKAEVQMKFKLVTPQDYARFKKSM 947  
QY 148 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPPEBGSAPRSWKGMNS 207  
Db 948 VKGLDSYEKEDKVIKEMAAQIREVEQSRQEVRSVLE-----TGPPR----- 990  
QY 208 QVASSLOQPSNLDLPPAPELDMWETGPSLTFIGHQDIPGVGNIHSGATPPPMIODEEYIA 267  
Db 991 -YALTVRSPAVLSR-----RTLKSGAPPPQ--TPEAHPQ 1021  
QY 268 GNOETGPSYEBFLKEKEQKLIK-KLPDPVRGANGFHDHSSRTSAGWLPSPGPRLEXWTPPLAV 326  
Db 1022 ARCLCAPRRGALKPEPPGRTKLGVPPTHTRKARPHAAKTSP-----RPRCTROAP--- 1072  
QY 327 QTSTPKLKL---QOXRSSHIOKKA 347  
Db 1073 -NKTOSLQLAGKARKTALHLQTKA 1095  
RESULT 4  
US-09-746-783-206  
; Sequence 206, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Fechtel, Kim  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

```

;
; ENCODING THEM
;
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/746,783
; FILING DATE: 21-Dec-2000
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Milasincic, Debra J.
; REGISTRATION NUMBER: 46,931
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 104 amino acids
; TYPE: amino acid
; STRANDEDNESS: <unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-746-783-206

Query Match          9.0%; Score 198; DB 10; Length 104;
Best Local Similarity 62.5%; Pred. No. 7.4e-09;
Matches 40; Conservative 5; Mismatches 15; Indels 4; Gaps 1;

Qy 201 SWKGMNSQVASSLQQPSNLDLPAPELDMWTGSLTFIGHQDIPGVGNHSGATPPWMI 260
Db : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19 AWMFFSQVASSLQQPSNLDLPAPELDMWTGSLTFIGHQ-----VQISKPEGQWSL 74

Qy 261 QDEE 264
Db : : :
75 RSQD 78

RESULT 5
US-10-424-599-209186
; Sequence 209186, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 209186
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(389)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_30923C.1.pep
US-10-424-599-209186

```

```

Query Match          6.8%; Score 148; DB 15; Length 389;
Best Local Similarity 21.4%; Pred. No. 0.0008;
Matches 60; Conservative 38; Mismatches 105; Indels 78; Gaps 10;

Qy 79 WCLCCGCEVREHLSHGNTLVLYGGLLEHLASPEHKKATNKFWENKAEVQMKFELYTPQ 138
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
90 WCVFCD-----QHDELNSSFACANAIHRLASAEHVKNLKQFFWKYGGAAQDLDAFVMSDD 145

Qy 139 DYARFKKSMVKGLDSEYEKEDKVIKEMAAQIREVEQSRQEVRSVLEPQAVPDPEEGSSA 198
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
146 DVDKWEKCTARKDEASEGSRRTVIGPSSDXNEQVS-----CSGLCGVATGATS 195

Qy 199 PRSWKGMNSQVASSLQQ-----PSNLDLP-----PAPELDMWTGSLTFI 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 SEVCSGANN---FSLQDFAVGRSSLSLPHDGRQRSSNGYSCNKKVRENGRMVSGES---- 248

Qy 240 GHQDIPGVGNHSGAT-----PMM-IODEYIAGNQEIGPSVEEFLKEKEKQK 289
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
249 SHQGVQMTQISTGSAGGVFSEPPPLGTDTDEVQMRSSNKLK-----KSK 295

Qy 290 KLPPDRVGNFDSH-----SRTSAGWLPSPG 315
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 KLSKRVGAANVVEKKIEMEKERRGESIRNECDSNWLPNFG 336

RESULT 6
US-10-425-115-306861
; Sequence 306861, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 306861
; LENGTH: 1040
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1040)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_42928C.1.pep
US-10-425-115-306861

Query Match          5.4%; Score 117.5; DB 16; Length 1040;
Best Local Similarity 22.7%; Pred. No. 1.12;
Matches 75; Conservative 48; Mismatches 105; Indels 103; Gaps 19;

Qy 4 TRWRSRIV---DPRAAXNSIRAHEGPPFCGRGCTFTTAASQRLKEAFERLLPQVEAARK 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
513 TRWSKAILMFEDDER-----FKAVERPREREDLFENYLVELHKKEK 553

Qy 61 AIRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNTLVLYGGLLEHL-----ASPEHKKAT 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
554 AKAAEEHKRYVAE-----YFAFLESCDFIKASTQWRKVQ 587

Qy 117 NKFWMENKAEVQMKFVLVTPQDYARFKSMVKGLDSYEKEDKVIKEMAAQIREVQSR 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
588 ERL--EDDERYSRLKPF-----DRLLDIFQYIRHLEKEEERQKRVQKD---QVRRQERKN 637

Qy 177 QEVVRSVLEPQAVPDPEEGS--SAPRSWKGMNSQVASS---LQPSNLDLPAPEL--DWM 230
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
638 RDGFRKMLEHVA-----DGTLNARTRDRYCAQIKDSQSYLAVASNTSGSTPKELFDVVI 693

```

```
QY 231 ETGSLTFIGHQ-----DIPGVNHSIHSATP---PMMIQ-----DDEYIAGNOEIGP 274
Db 694 EE-----LGKQYQEDTQIKEV--VKSGKIPMTTSWTLEEFOTATLEDDALKGISTINI 745
QY 275 S--YE---EFLKEKEKQKLKLPDRVGANF 300
Db 746 KLVYDDQLRLKKEQKEAKK--RORLGENF 774

RESULT 7
US-10-425-114-41619
; Sequence 41619, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 41619
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3078-055-D7_FLI.pep
US-10-425-114-41619
```

```
Query Match 5.3%; Score 116; DB 15; Length 367;
Best Local Similarity 20.5%; Pred. No. 0.39;
Matches 61; Conservative 43; Mismatches 92; Indels 102; Gaps 12;

QY 4 TRWRSRIV---DPRAAXNSIRAHEGPFPCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
Db 17 TRWSKAILMFEDDER-----FKAVERPREREDLFENYLVHLHKEK 57
QY 61 ATRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNLTVLVYGLLEHL-----ASPEHKAT 116
Db 58 AKAAEHKRYVAE-----YRAFLESCDFIKASTQWRKVQ 91
QY 117 NKFWENKAEVQMKKFLVTPDYARFKSMVKGLDSYEKKEDVKIKEMAAQIREVEQSR 176
Db 92 ERL--EDDERYSRLKFX-----DRLDIFOEYIRHLEKEEEOQKRVQKD---QVRQERKN 141
QY 177 QEVRSVLEPQAVDPPEGS--SAPRSWKGMNSQVASSLQQPSNLDLPPAPELDMMETGFS 235
Db 142 RDGFRKMLEEHA-----DGTLNARTWRDYCAQIKDS-----174
QY 236 LTFIGHQDIPGVNHSIHSATPMTQDEYIAGNOEIGPSY--EEFLKEKEKQKLKLP 292
Db 175 -----QSLAVASNTSGTPELFDDV-----IEELGKQYQEDKIQIKEVVKSGLP 221
```

```
RESULT 8
US-10-425-114-69227
; Sequence 69227, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69227
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3076-041-ELL_FLI.pep
US-10-425-114-69227
```

```
Query Match 5.3%; Score 116; DB 15; Length 369;
Best Local Similarity 20.5%; Pred. No. 0.39;
Matches 61; Conservative 43; Mismatches 92; Indels 102; Gaps 12;

QY 4 TRWRSRIV---DPRAAXNSIRAHEGPFPCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
Db 18 TRWSKAILMFEDDER-----FKAVERPREREDLFENYLVHLHKEK 58
QY 61 ATRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNLTVLVYGLLEHL-----ASPEHKAT 116
Db 59 AKAAEHKRYVAE-----YRAFLESCDFIKASTQWRKVQ 92
QY 117 NKFWENKAEVQMKKFLVTPDYARFKSMVKGLDSYEKKEDVKIKEMAAQIREVEQSR 176
Db 93 ERL--EDDERYSRLKFX-----DRLDIFOEYIRHLEKEEEOQKRVQKD---QVRQERKN 142
QY 177 QEVRSVLEPQAVDPPEGS--SAPRSWKGMNSQVASSLQQPSNLDLPPAPELDMMETGFS 235
Db 143 RDGFRKMLEEHA-----DGTLNARTWRDYCAQIKDS-----175
QY 236 LTFIGHQDIPGVNHSIHSATPMTQDEYIAGNOEIGPSY--EEFLKEKEKQKLKLP 292
Db 176 -----QSLAVASNTSGTPELFDDV-----IEELGKQYQEDKIQIKEVVKSGLP 222
```

```
RESULT 9
US-10-425-114-51821
; Sequence 51821, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51821
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700100934_FLI.pep
US-10-425-114-51821
```

```
Query Match 5.3%; Score 116; DB 15; Length 653;
Best Local Similarity 20.5%; Pred. No. 0.85;
Matches 61; Conservative 43; Mismatches 92; Indels 102; Gaps 12;

QY 4 TRWRSRIV---DPRAAXNSIRAHEGPFPCGRGTFTAASTORQLKEAFERLLPQVEAARK 60
Db 303 TRWSKAILMFEDDER-----FKAVERPREREDLFENYLVHLHKEK 343
QY 61 ATRAAQVERVYVPEHERCCWCLCCGCEVREHLSHGNLTVLVYGLLEHL-----ASPEHKAT 116
Db 344 AKAAEHKRYVAE-----YRAFLESCDFIKASTQWRKVQ 377
```



```
Qy 117 NKFWENKAEVOMKEKFLVTPQDYARFKKSMVGLDSVEEKEDKVIKEMAAQIREVQSR 176
Db 378 ERL--EDDERYSRLEKF-----DRLDIFQEYIRHLEKEEQRVKQD---QVRRQERKN 427
Qy 177 QEVRSVLEPQAVDPDEGS-SAPRSWKGMNSQVASSLQQPSNLDLPPAPELDMMETGPS 235
Db 428 RDGRFKMLEHVA-----DGLNATRWRDYCAQIKDS----- 460
Qy 236 LTFIGHQDIPGVNIHSGATPPMMIQDEEYIAGNQEIGPSY-EBFLKEKEKQKLKLP 292
Db 461 -----QSLAVASNTSGSTKPLFDDV-----IEELGKQOEDKIQIKVEVKGKIP 507

RESULT 10
US-10-425-115-219898
; Sequence 219898, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 219898
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_132134C.1.pap
US-10-425-115-219898
```

```
Query Match 5.1%; Score 112; DB 16; Length 160;
Best Local Similarity 28.3%; Pred. No. 0.27;
Matches 34; Conservative 17; Mismatches 39; Indels 30; Gaps 5;

Qy 249 NIHSGATPPMMIQDEEYIAGNQEIGPSVEEFLKEKEKQKLKLPDPRVGANFDHSSR--- 305
Db 22 NVHTGAPPPMLKANE-----HDPKNLSKSCGLSSRGKLRKLPNPNRVGAWAERRRAEM 76
Qy 306 -----TSAGWLPSPGPRLEXWTPPLAVTSTPKLKLOQXRS--IQKANHAL 351
Db 77 EMEKQGEIVPATSDSSWLPNFG---SVW-----QSGTRKESRKEFEKSHKIHDTKSDHDL 128
```

```
RESULT 11
US-10-425-114-62537
; Sequence 62537, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62537
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700349560_FLI.pap
US-10-425-114-62537
```

```
Query Match 5.1%; Score 112; DB 15; Length 253;
Best Local Similarity 28.3%; Pred. No. 0.51;
Matches 34; Conservative 17; Mismatches 39; Indels 30; Gaps 5;

Qy 249 NIHSGATPPMMIQDEEYIAGNQEIGPSVEEFLKEKEKQKLKLPDPRVGANFDHSSR--- 305
Db 115 NVHTGAPPPMLKANE-----HDPKNLSKSCGLSSRGKLRKLPNPNRVGAWAERRRAEM 169
Qy 306 -----TSAGWLPSPGPRLEXWTPPLAVTSTPKLKLOQXRS--IQKANHAL 351
Db 170 EMEKQGEIVPATSDSSWLPNFG---SVW-----QSGTRKESRKEFEKSHKIHDTKSDHDL 221
```

```
RESULT 12
US-10-425-114-67550
; Sequence 67550, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67550
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-078-C12_FLI.pap
US-10-425-114-67550
```

```
Query Match 5.1%; Score 112; DB 15; Length 253;
Best Local Similarity 28.3%; Pred. No. 0.51;
Matches 34; Conservative 17; Mismatches 39; Indels 30; Gaps 5;

Qy 249 NIHSGATPPMMIQDEEYIAGNQEIGPSVEEFLKEKEKQKLKLPDPRVGANFDHSSR--- 305
Db 115 NVHTGAPPPMLKANE-----HDPKNLSKSCGLSSRGKLRKLPNPNRVGAWAERRRAEM 169
Qy 306 -----TSAGWLPSPGPRLEXWTPPLAVTSTPKLKLOQXRS--IQKANHAL 351
Db 170 EMEKQGEIVPATSDSSWLPNFG---SVW-----QSGTRKESRKEFEKSHKIHDTKSDHDL 221
```

```
RESULT 13
US-10-607-631-20
; Sequence 20, Application US/10607631
; Publication No. US20040091901A1
; GENERAL INFORMATION:
; APPLICANT: Minion, F. Chris
; APPLICANT: Djordjevic, Steven P.
; TITLE OF INVENTION: Immunogenic Mycoplasma hyopneumoniae
; FILE REFERENCE: 08411/035001
; CURRENT APPLICATION NUMBER: US/10/607,631
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392,632
; PRIOR FILING DATE: 2002-06-28
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1878
; TYPE: PRT
; ORGANISM: Mycoplasma hyopneumoniae
US-10-607-631-20
```

```
Query Match          5.0%; Score 110.5; DB 15; Length 1878;
Best Local Similarity 23.7%; Pred. No. 11;
Matches 71; Conservative 45; Mismatches 132; Indels 51; Gaps 14;

QY 92 SHGNLTVLYGGLLEHL-----ASPEHKATNKFWENKAEVQMKKFL--VTPQDYA 141
DB 414 SLGYNFLPDDLASHLDYTLVSKAKIKQSSITKKLFIELPIKISLSSILGDOENIKT 473

QY 142 RFKKSVMVGLDSYEEKE-DKVIKEMAAQI-REVEQSRQEVVRSVLEPQAVPDEGS-SA 198
DB 474 LFEKVTFLKDNFRDVEIEKAFGLLYPGVNEELEQARREQASLEKEKAKGLKFSQOK 533

QY 199 PRSWKGMNSOVASSLOOPSND--LPPAPELDWMTGPSITFGHODIPGVGNHSGATP 256
DB 534 DENLKAINQ--DGLREDDNITELPENSPIQ-----QOEKAGLG---SSPK 577

QY 257 PWMIOBEE----YIAGNQEIGPSYEEFLKEKEKQKLKLPDRVGANFDHSSRTSAGMLP 312
DB 578 PYMIKDVQNRYYLAKSQ-----IQELIKAKDYTKLAKLSNRHTYNSLRLEKQ---LF 629

QY 313 SFGPRLXWTPLA----VQTSTPKLKLQXRRS-----HIQKANHALYQLPXGXPKST 363
DB 630 EVNPRIPSSRDIEAKFVLDKTEKNYQIYSSASPAFQNKWSLFGYYRYLLGLDPKQT 688

RESULT 14
US-10-437-963-133129
; Sequence 56, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133129
; LENGTH: 1109
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_3502C.1.pep
US-10-437-963-133129

Query Match          5.0%; Score 109.5; DB 16; Length 1109;
Best Local Similarity 22.2%; Pred. No. 6.3;
Matches 76; Conservative 43; Mismatches 126; Indels 97; Gaps 18;

QY 90 HLSHGNLTVLYGGLLEHLASPEHK-KATNKFWEN-----KAEVQMKKFL---- 134
DB 8 HAEKKLNLLQCNQDKVAFATHQLQGPASAWDNHMATRPDDVTRDAEKQKFMAGLD 67

QY 135 --VTQ-----DYARPKSMVGLDSYEEKEKDKVIKEMAAQIREVEQSRQEVVRSVLEPQA 188
DB 68 DELTNQLISGDYADPERLVDAIROEDQNRKMDRKRKAAQFRAHQGSHQ-----RPRF 120

QY 189 VPDPEGSSA--PRSWKGMNSOVASSLOOPSNDLPPAPELDWMTGPSITFGHODIPG 246
DB 121 TPGQGGGPTIMVIRYRFN-----PSN-----FHQAS-----GSQDQH 156

QY 247 VGNTHSGATPMMIODEYIAGNQIGPSYEEFLKEKEKQKLKLPDRVGANFDHSSRT 306
DB 157 -GQLNRGAAP-----RPPVAPAQSGPS-----AQAKKETGKPGSCFCNG--- 195
```

```
QY 307 SAGWLPSTGPRLEXWTPPLAVQTSTPKLKLQXRRSHIOKKANHALYQLPXGXPKSTKPL 366
DB 196 ELGHADLADCKPKPRRAGP-----RFVQARVNHASAEBAQAAPVVLGTFFVNSIPA 245

QY 367 LYLPPKFFIIVFLRQTYSTFXFNKVLFPGLRGL-ESSLSIP 407
DB 246 -----TVLFDSGATHSFISKK---FVGMHGLVREELSTP 276

RESULT 15
US-10-051-874-56
; Sequence 56, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Saeha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Macdougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol BA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennnda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
```

; PRIOR APPLICATION NUMBER: 60/261,376  
; PRIOR FILING DATE: 2001-01-16  
; NUMBER OF SEQ ID NOS: 269  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 56  
; LENGTH: 4952  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-051-874-56

Query Match 5.0%; Score 109; DB 15; Length 4952;  
Best Local Similarity 20.9%; Pred. No. 54;  
Matches 80; Conservative 51; Mismatches 149; Indels 102; Gaps 15;

Qy	101	GGLLEHLASPE-----HKKATNKFWMENKA-----EVQMKKEFLVTPQDYAR	142
Db	2609	GGPPAHLITPSPLSGPGGSSLEKFELESALTLPGGPAASGDELDMESSLVASE----	2664
Qy	143	FKSMVKGLDSYEKEDKVIKEMAAQIREVEQSREQEVVRSVLEP-----QAVDPPEGSS	197
Db	2665	-LPLLLIEDLLEHEKELQKKQLSAQLQPAQQQQQQQQQQHSLLPAPGPAQAMSLPHEGSS	2723
Qy	198	APRSWKGMNSOVA-----SSLQPSNLDPAPAPLDMETGPSLTFIGHQ-----	242
Db	2724	P--SLAGSQQLSLGLAVARQFGLPQPLMPTQPPAHALQ-ORLAPSMAMVSNQGHMLSGQ	2780
Qy	243	-----DIPGVNIHSGATPPMMIQDEEYIAGNQIGPSY-----EEF	279
Db	2781	HGGQAGLVPOQSSQPVLSQKPMGTMPSPCMKPPQQLAMQQQLANSFFPDTDLDKFAAEDI	2840
Qy	280	LKEKEKQK-----LKKLPDRVGNANFDHSSRTSAGWLPSPGPRLEXWTPPLAVQTSTPKL	333
Db	2841	IGPIAKAMVALKGIKKV-----MAQGSIGVAPGMNRQ-----QVSLLAQ	2880
Qy	334	KLQQRSSHIOKXANHALYQLPGXKPKSTKPLLVLPKPF--IIVFLRKOTYSFIKFN-	390
Db	2881	RLSGPSSDLQNHVAAGSGQERSAGDSQPRP---NPPTFAQGVINEADQRYEENLFT	2937
Qy	391	-KVLFFGLRGLLESSLIPSTSR	411
Db	2938	QQLLMQLKVLEEQIGVHRKSR	2959

Search completed: June 7, 2005, 12:25:33  
Job time : 54.7607 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:51:20 ; Search time 16.3155 Seconds  
(without alignments)  
1907.914 Million cell updates/sec

Title: US-09-155-676B-5  
Perfect score: 2190  
Sequence: 1 IGVTWRNRRIKVPRAAXNS.....RGLESSISPTSGRGTRHF 417

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued patents AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/PTUS.COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108.5	5.0	1225	4	US-09-501-171-4
2	108.5	5.0	1225	4	US-09-949-016-6063
3	107.5	4.9	1540	4	US-09-949-016-7037
4	104	4.7	1525	4	US-09-418-710-69
5	104	4.7	1525	4	US-09-839-479-68
6	103.5	4.7	578	3	US-09-066-046-6
7	103.5	4.7	753	4	US-09-949-016-6676
8	103.5	4.7	3224	2	US-08-705-660-34
9	103.5	4.7	3224	3	US-08-989-045-34
10	103.5	4.7	3224	4	US-09-538-092-1161
11	103.5	4.7	3224	4	US-09-315-355A-34
12	103	4.7	1527	4	US-09-418-710-27
13	103	4.7	1527	4	US-09-839-479-27
14	103	4.7	1531	4	US-09-418-710-29
15	103	4.7	1531	4	US-09-839-479-29
16	102.5	4.7	667	4	US-09-949-016-7759
17	102.5	4.7	777	4	US-09-949-016-8706
18	102.5	4.7	777	4	US-09-949-016-8707
19	102.5	4.7	1009	2	US-08-357-642A-1
20	102.5	4.7	1009	2	US-08-460-626-1
21	102.5	4.7	1014	4	US-09-949-016-11533
22	101	4.6	2101	1	US-08-456-390-4
23	101	4.6	2101	1	US-08-470-950-4
24	101	4.6	2101	1	US-08-467-781-4
25	101	4.6	2101	1	US-08-195-487-4
26	101	4.6	2101	2	US-08-483-924-4
27	101	4.6	2101	3	US-09-452-294-1
28	101	4.6	2101	3	US-09-452-294-1
29	101	4.6	2101	3	US-09-452-294-1
30	101	4.6	2101	3	US-09-452-294-1
31	101	4.6	2101	3	US-09-452-294-1
32	101	4.6	2101	3	US-09-452-294-1
33	101	4.6	2101	3	US-09-452-294-1
34	101	4.6	2101	3	US-09-452-294-1
35	101	4.6	2101	3	US-09-452-294-1
36	101	4.6	2101	3	US-09-452-294-1
37	101	4.6	2101	3	US-09-452-294-1
38	101	4.6	2101	3	US-09-452-294-1
39	101	4.6	2101	3	US-09-452-294-1
40	101	4.6	2101	3	US-09-452-294-1
41	101	4.6	2101	3	US-09-452-294-1
42	101	4.6	2101	3	US-09-452-294-1
43	101	4.6	2101	3	US-09-452-294-1
44	101	4.6	2101	3	US-09-452-294-1
45	101	4.6	2101	3	US-09-452-294-1

28	101	4.6	2101	5	PCT-US93-06160-4	Sequence 4, Appli
29	101	4.6	2107	4	US-09-949-016-7646	Sequence 7646, Ap
30	101	4.6	2107	4	US-09-949-016-7647	Sequence 7647, Ap
31	100.5	4.6	578	3	US-08-975-762-50	Sequence 50, Appl
32	100.5	4.6	578	3	US-09-295-028-50	Sequence 50, Appl
33	100.5	4.6	578	3	US-09-106-582-50	Sequence 50, Appl
34	100.5	4.6	578	4	US-09-159-469-50	Sequence 50, Appl
35	100.5	4.6	578	4	US-09-693-542-50	Sequence 50, Appl
36	100	4.6	2907	4	US-09-698-295-1	Sequence 1, Appli
37	99	4.5	475	4	US-09-949-016-6317	Sequence 6317, Ap
38	99	4.5	487	4	US-09-949-016-11205	Sequence 11205, A
39	99	4.5	1805	1	US-07-853-913-2	Sequence 2, Appli
40	98.5	4.5	702	4	US-09-248-796A-15814	Sequence 15814, A
41	98.5	4.5	8991	4	US-08-714-741-32	Sequence 32, Appl
42	97	4.4	305	3	US-08-937-271-10	Sequence 10, Appl
43	97	4.4	739	4	US-09-949-016-9709	Sequence 9709, Ap
44	95.5	4.4	273	2	US-08-592-214A-8	Sequence 8, Appli
45	95.5	4.4	273	3	US-08-659-188-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-09-501-171-4  
; Sequence 4, Application US/09501171  
; Patent No. 6783982  
; GENERAL INFORMATION:  
; APPLICANT: St. George-Hyeloop, Peter H.  
; APPLICANT: Fraser, Paul E.  
; TITLE OF INVENTION: Proteins Related to Neuronal  
; FILE OF INVENTION: Regeneration and Uses Thereof  
; FILE REFERENCE: 1034/1F811  
; CURRENT APPLICATION NUMBER: US/09/501.171  
; CURRENT FILING DATE: 2000-02-09  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 1225  
; TYPE: PRT  
; ORGANISM: human  
US-09-501-171-4

Query Match	5.0%	Score 108.5;	DB 4;	Length 1225;
Best Local Similarity	22.7%	Pred. No. 0.41;		
Matches	81;	Conservative	46;	Mismatches 111; Indels 119; Gaps 19;
Qy	31	GRGTTAASTQRQLKE---	AFERLLPQVEARAKAIRAAQVRYVPEHERCCWCLCCGCEV	87
Db	42	GSETTTTSAIASYKEQELQFERLTREAEQIV-ASQLER-----	CKL	85
Qy	88	REHLSHGNLTVLYGGLLEHLASPEHKATNKFWMENKAEVQMKEFLVTPQDYARFKSM	147	
Db	86	-----GSET-----GSMSSMSAE-----EQFMQSQ-----	107	
Qy	148	VKGLDSYEEKEDKV---IKEMAAQIREVQSRQVRSVLEPQAVPDPPEGSSAPRSWK	204	
Db	108	-----DGQKDIEDELTTCGLVDSCLRSQES-----	GILDPQDYSTGERSLLSQSALQ	157
Qy	205	MNSQVASSLOQP-----SNLDLP-----	PAPELDMETGPSL-----TFIGHQDIPG	246
Db	158	LNSKPEGSFYPAYSYHSNQTALGETTPSQLPARGTQARATGQSFGQTTSRAGHLGP-	216	
Qy	247	VGNHSGATPMWLODEYTAGNQEIGPSYEEFLKEKEKQKLKLPDRVCANFDHSRT	306	
Db	217	-----EPAPPPPPPPPPPPFA---PSLGSAPH-----	LPDAPPAANAALYYSST	258
Qy	307	SAGWLPSPGRLXWTPLAVQTSPTKLQOXRSSHIQKXANHALYOLPXGXKPKST	363	
Db	259	-----LPA-PPR-----GGSLAAPQGSPTKLQRGSA-----	PEGATYAAPRGSSPKQS	303

RESULT 2

US-09-949-016-6063  
; Sequence 6063, Application US/09949016  
; Patent No. 6812339  
; ORGANISM: Human  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6063  
; LENGTH: 1225  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6063

Query Match 5.0%; Score 108.5; DB 4; Length 1225;  
Best Local Similarity 22.7%; Pred. No. 0.41;  
Matches 81; Conservative 46; Mismatches 111; Indels 119; Gaps 19;

QY 31 GRTFTTAASQOROLKE---AFERLLPQVEAAKAIKRAAQAQVYVPEHERCCWCLCCGCEV 87  
DB 42 GSETTTSAILASVKQEQLQEFELTLEAEQIV-ASQLER-----CVL 85

QY 88 REHLSHGNLTVLYGGLLEHSLASPEHKATNKPWENKAQVOMKEKFLVTPQDYARFKSM 147  
DB 86 -----GSET-----GSMSSMSAE-----EQFOWQSQ----- 107

QY 148 VKGLDSYEKEDKV---IKEMAAQIREVQSRQEVVRSVLEQVAPDPEEGSAPRSWK 204  
DB 108 -----DGOKDELTGTELVDSICRSLOES-----GLDPDYTGTPGRLSLSALQ 157

QY 205 MNSQVASSLQOOP-----SNLDLP-----PAPELDWMTGSL-----TFIGHQDIPG 246  
DB 158 LNSKPEGSFYPPASYHSNQTALGETTPSQLPARGTQARATGQSFSGTTSRAGHLA 216

QY 247 VGNHSGATPPWMIQDEYTAGNQEGPSYEFLKEKEKOKLKLPPDRVGFANFHDHSSRT 306  
DB 217 -----EPAPPPPPPPPEPFA---PSLGSAPF-----LPDAPPAALAAALYSSST 258

QY 307 SAGHLPSPGRLXETWTPAVQTSPTPKLKLOQXRSHHQKKAHALYQLPXGKXPKST 363  
DB 259 -----LPA-PPR---GGSPLAAPQGSPTKLQGGSA-----PEGNTYAAPRGSSPKQS 303

RESULT 3  
US-09-949-016-7037  
; Sequence 7037, Application US/09949016  
; Patent No. 6812339  
; ORGANISM: Human  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7037

; LENGTH: 1540  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7037

Query Match 4.9%; Score 107.5; DB 4; Length 1540;  
Best Local Similarity 20.9%; Pred. No. 0.72;  
Matches 104; Conservative 60; Mismatches 126; Indels 207; Gaps 30;

QY 14 PRAAXNSIRAHGPPFCGRG-----TFTAASQOROL---KEAFERLLP-----QVEA 57  
DB 26 PPGAWRSRA-AGPSAAPRGPQPPVPTAAAMAPLLGRKFPPLVKPLPGSEPLTIPH 84

QY 58 ARKAIKRA-----AQVERVYVPEHERCCWCLCCGCEVREHLSHGNLTVLYGGLLEHSLASPEH 112  
DB 85 TQEAFTREYEYERLERY---SERIWTCKSTG-----SSQLT-----H 119

QY 113 KKATNKPWENKAQVQ---MKEKFLVTPQDYARFKSMV-----KGLDS----- 153  
DB 120 KEA-----WEEEQEVAELLKEEF---PAWYEKLIVLEMVHHINTASLEKLVDTAWLEIMTKY 171

QY 154 -----YEEKEDKVIKEMAAQIREVQSRQEVVRSVLEQVAPDPEEGS-SAPRSWKGM 205  
DB 172 AVGEEDCFVVGKMKMLKIVKIHPLEK-----VDEEATEKKSDGACDSSDKEN 222

QY 206 NSQVASSLQO-----PSNLD-----LP----- 222  
DB 223 SSQIAQDHQKKEIVVKEDEGRRESINDRARRSPKLPSTSLKGERKWAPPKFLPHKYDVK 282

QY 223 -----PAPELDWMTGSLT-----FIGHQDIPGVGNHSGATPPWMIQDEEYIA 267  
DB 283 LQNEDKIISNVPADSLIRTERPPNKEIVRYFIHNAU-----RAGTGENAPWVEDE---LV 336

QY 268 GNQIGSPSYEEFLKEKEK-----OKLKLKLPDDR---VGANFHDHSSRTSAGWLPSPFGP 316  
DB 337 KKYSLPSKFSDFLLDPKYMTLNPSTKRKNTGSDPRKPSKSKTDNSSLSS-----PLNP 391

QY 317 RLEKW-----TPLAVQTS-----TPKCLKLOQ-----XRSSHIQKKAHAL 351  
DB 392 KL---WCHVHLKKSLSGSPKLVKNSKNSKSPPEEHEEMKQMSPNKLTNTHFIPKKG--- 445

QY 352 YQLPXGKXP---KSTKPL 366  
DB 446 ---PPAKPKGHSDKPL 459

RESULT 4  
US-09-418-710-69  
; Sequence 69, Application US/09418710  
; Patent No. 6596482  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Michael H.  
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR  
; FILE REFERENCE: 06501-042001  
; CURRENT APPLICATION NUMBER: US/09/418,710  
; CURRENT FILING DATE: 1999-10-15  
; PRIOR APPLICATION NUMBER: PCT/JF98/01783  
; PRIOR FILING DATE: 1998-04-17  
; PRIOR APPLICATION NUMBER: JP 9/310027  
; PRIOR FILING DATE: 1997-10-24  
; PRIOR APPLICATION NUMBER: JP 9/116570  
; PRIOR FILING DATE: 1997-04-18  
; NUMBER OF SEQ ID NOS: 73  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69  
; LENGTH: 1525  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-418-710-69

Query Match 4.7%; Score 104; DB 4; Length 1525;  
Best Local Similarity 21.3%; Pred. No. 1.5;  
Matches 99; Conservative 53; Mismatches 109; Indels 204; Gaps 29;

Query Match	4.7%	Score 103.5;	DB 3;	Length 578;
Best Local Similarity	20.8%	Pred. No. 0.4;		

Matches	64; Conservative	51; Mismatches	114; Indels	79; Gaps	13;
Qy	13	DPRAAXNSIRAHEGPPFCRGFTTAASTQRQKEAFERLLPQVEAARAKAIRAAQVRYV-	71		
Db	10	DYRIGSETISSGDDGYEGCACDDASTNAYSD-----KCRVVGRTWRPELVLYVG	62		
Qy	72	PEHERCCWCLCCGCEVREHLHSGNLT-----VLYGGLLEHLASPEHK	113		
Db	63	DEHVAC-----RDVASGMHGNLPKGVYFIABAGRAATAEGGVTTTVEALSIVQEE	115		
Qy	114	KATNKFWENKAEVQMKFLV-----TPQ--DYARFKSMVGLDSEYKEKDK	160		
Db	116	EGTGMYL-----INAEKAVVRFFKIEKSAABEPQTVPSPVESATGSGVDTEEQE--	167		
Qy	161	VTKEMAAQIREVQSRQSVLRSLVPQVPPESGSS-----APRSWKGNNS--QVASSL	213		
Db	168	-IDQEPALPEETEEQEV--IIEEGTLIDLEQFVQVPVVAEALPGVEAAEATVPSL	223		
Qy	214	QQPSNLIDLPPAPELDMWETGPSLTF-----IGHQIDIPGVNTHSGATPMMWIQDB	263		
Db	224	EENKLEQVVVAPEAQQLSAPEVSPAQPESTVLGVAEGLDKSEVSVEANADVP-----QK	279		
Qy	264	EYIAGNQE	271		
Db	280	EVISGQOE	287		

```

RESULT 7
US-09-949-016-6676
; Sequence 6676, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6676
; LENGTH: 753
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6676

```

Query Match	4.7%;	Score 103.5;	DB 4;	Length 753;
Best Local Similarity	22.0%;	Pred.No. 0.59;		
Matches	84;	Conservative	48;	Mismatches 149; Indels 101; Gaps 19;
QY	19	NSIRAH-----EGPFFCG-RGTFTAASTQRQLK-----BAFERLLRQVE	56	
Db	185	HSLRIHVVRVHTKEKPFECDDVOGCEKAFNTLYRLKAHQRLHGTGFNCESGCCKFYTTL	244	
QY	57	AARKATRAAQR-----YPEH-----ERCWCCLCGCCE-----	86	
Db	245	DLRKHILRTHTGEKPFRCDDHCCKAKFAASHHLKTHVRTHTGERFFCFSGNCEKTFTSTQY	304	
QY	87	-VREHL--SHGNLTAVLYGGILLEHLASPHKK-----ATNKFWWENKAEOVMQMKFL	134	
Db	305	SLKSHMKGHDNGHSYNALPQHNGSEDTNHSLCULSDLSTDSLENSNSTTQQDLSLST	364	
QY	135	VTPQDYARPKSMVKGLDSYEEDKVIKEMAAQIREVQSQRVETSVULEPQAVPDPEE	194	
Db	365	ISP---AIIFES-----FQNSDDTAIQBDPOQTASLTFSFGDAESVD---VP-PST	411	
QY	195	GSSAPRS-----WKGMNSQVASSLOQPSNLPLPAELDMWTMGPSITFIHQDPGVGN	249	

```

Db      412  GNSASLSPLVLQPLGLSEPPQLLPASAFSAPPPAPSL-----GP-----GSQQ-AAFGN 460
Qy      250  IHSGATPPMIODEEY- IAGNQEIIGSYSEFLKEKEKQKLKLPDPDVGANTDHSRRTSA 308
Db      461  -----PPALQPPPEVPHSTQFAANHOEFLPHQPAQ-PIVPGLSVVAGASASAAVA 513
Qy      309  GWLPSFGPRLEXWTPLAVQTST 330
Db      514  SAVAAPAPQSTTEPLPAMVQT 535

RESULT 8
US-08-705-660-34
; Sequence 34, Application US/08705660
; Patent No. 5858683
; GENERAL INFORMATION:
; APPLICANT: KEESEE, SUSAN
; APPLICANT: OBAR, ROBERT
; APPLICANT: WU, YING-JYE
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault
; STREET: 125 High St.
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,660
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: GREENHALGH, DUNCAN A
; REGISTRATION NUMBER: 38,678
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3224 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-705-660-34

```

QY 195 GSSAPRS-----WKGMSQVASSLQOPSNLDLPPAPELDWMTGPSLFTTGHQDIPGVGN 249



Db 758 EGGPLYKNGSLRNADSEIKRSTPSPTRYSLSPSKSYK-----Y 795  
Qy 252 SGATPPMMIODEEYIAGNQEIGPSYEFLEKEKEKQL-----KKLPPDRVGA 298  
Db 796 SPKTPPRAEDQNSLL--KMICQVEAIKKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853  
Qy 299 NFDHSSSTSAGWLPSPGFRLEXWTPLAVQSTPKLKLQQRSSHIQKANHLYQLPXGX 358  
Db 854 GY-QGSQTFHG-----APLTVATTGSPV-----YYSQSPAYNSQYLLRPAA 893  
Qy 359 KPKSTKPLLY---LPPKFFIIVFLRK-----QTSYFIKFNKVL 394  
Db 894 NVTPTKGPVYGMNRLPPQOHYIAYPQQMHTPPVQSSACMFSEQMY 939  
RESULT 9  
US-08-989-045-34  
; Sequence 34, Application US/08989045  
; Patent No. 6027905  
; GENERAL INFORMATION:  
; APPLICANT: KEESEE, SUSAN  
; APPLICANT: OBAR, ROBERT  
; APPLICANT: WU, YING-JYE  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
; DETECTION OF CERVICAL CANCER  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Testa, Hurwitz & Thibault  
; STREET: 125 High St.  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/989,045  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GREENHALGH, DUNCAN A  
; REGISTRATION NUMBER: 38,678  
; REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3224 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-989-045-34  
Query Match 4.7%; Score 103.5; DB 3; Length 3224;  
Best Local Similarity 18.2%; Pred. No. 5.3;  
Matches 74; Conservative 69; Mismatches 132; Indels 131; Gaps 17;  
Qy 48 FERLLPQVEAARKA-----IRAAQVERVYVPEHERCCWCLCGCEVRE 89  
Db 606 WKVPLPLLLKIIKKNSIPEIDPLFKHFHSDVIOASEIVEY-----BEDA 650  
Qy 90 HLSHGNLTVLYGGLLEHLASPEHKKATNKFW-----WENKA-----EVQMKEF 133  
Db 651 HITAILDAVNGNIEDAVTAFESIKSVSVYNNLALIFHRKAEDIENDALSPEEQECKNY 710  
Qy 134 LVTPQDYARFKKSMVKGDSYEEKEDKVIKEMAAQIREVQSRQSVRSVLEPQAVDPPE 193  
Db 711 LRKTRDY-----LIKIIDD--SDSNLSVVKLPVPLESV-----KEMLSNV-----OLEEDYS 757

Qy 194 EGSSAPR--SWKGNGNSQVASSLQOPSNLDPAPPELDWMETGPSLTFIGHODIPGVGNIH 251  
Db 758 EGGPLYKNGSLRNADSEIKRSTPSPTRYSLSPSKSYK-----Y 795  
Qy 252 SGATPPMMIODEEYIAGNQEIGPSYEFLEKEKEKQL-----KKLPPDRVGA 298  
Db 796 SPKTPPRAEDQNSLL--KMICQVEAIKKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853  
Qy 299 NFDHSSSTSAGWLPSPGFRLEXWTPLAVQSTPKLKLQQRSSHIQKANHLYQLPXGX 358  
Db 854 GY-QGSQTFHG-----APLTVATTGSPV-----YYSQSPAYNSQYLLRPAA 893  
Qy 359 KPKSTKPLLY---LPPKFFIIVFLRK-----QTSYFIKFNKVL 394  
Db 894 NVTPTKGPVYGMNRLPPQOHYIAYPQQMHTPPVQSSACMFSEQMY 939  
RESULT 10  
US-09-538-092-1161  
; Sequence 1161, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Giot, Loic  
; APPLICANT: Mansfield, Traci A.  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538,092  
; CURRENT FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CurapatSeqFormatter Version 0.9  
; SEQ ID NO 1161  
; LENGTH: 3224  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)...(0)  
; OTHER INFORMATION: Polypeptide Accession Number P49792  
US-09-538-092-1161  
Query Match 4.7%; Score 103.5; DB 4; Length 3224;  
Best Local Similarity 18.2%; Pred. No. 5.3;  
Matches 74; Conservative 69; Mismatches 132; Indels 131; Gaps 17;  
Qy 48 FERLLPQVEAARKA-----IRAAQVERVYVPEHERCCWCLCGCEVRE 89  
Db 606 WKVPLPLLLKIIKKNSIPEIDPLFKHFHSDVIOASEIVEY-----BEDA 650  
Qy 90 HLSHGNLTVLYGGLLEHLASPEHKKATNKFW-----WENKA-----EVQMKEF 133  
Db 651 HITAILDAVNGNIEDAVTAFESIKSVSVYNNLALIFHRKAEDIENDALSPEEQECKNY 710  
Qy 134 LVTPQDYARFKKSMVKGDSYEEKEDKVIKEMAAQIREVQSRQSVRSVLEPQAVDPPE 193  
Db 711 LRKTRDY-----LIKIIDD--SDSNLSVVKLPVPLESV-----KEMLSNV-----OLEEDYS 757  
Qy 194 EGSSAPR--SWKGNGNSQVASSLQOPSNLDPAPPELDWMETGPSLTFIGHODIPGVGNIH 251  
Db 758 EGGPLYKNGSLRNADSEIKRSTPSPTRYSLSPSKSYK-----Y 795  
Qy 252 SGATPPMMIODEEYIAGNQEIGPSYEFLEKEKEKQL-----KKLPPDRVGA 298  
Db 796 SPKTPPRAEDQNSLL--KMICQVEAIKKEMQELKLNSSNSASPHRWPTENYGPDSVPD 853  
Qy 299 NFDHSSSTSAGWLPSPGFRLEXWTPLAVQSTPKLKLQQRSSHIQKANHLYQLPXGX 358  
Db 854 GY-QGSQTFHG-----APLTVATTGSPV-----YYSQSPAYNSQYLLRPAA 893



[illegible]

Mon Jun 13 13:27:15 2005

```
QY 216 ----PSNLD-----LP-----PAPELDWMTGPSLT----FI 239
Db 198 PRKLPSTLKGGERKWAPPKFLPHKYDVKLQNEDKIISNVPADSLIRTERPPNKEIVRYFI 257
QY 240 GHQDIPGVGNIHSGATPPWMIQDEBEYIAGNQBIGPSVEEFLKEKEK-----QKLKKL 291
Db 258 RHNAL-----RAGTGENAPWVVEDE--LVKKYSLPSKFSDFLLDPYKYMTLNPSTKRKNTG 311
QY 292 PPDR---VCANFDHSSRTSAGWLPSPGPRLEW-----TPLAVQTS-----TPKL 333
Db 312 SPDRKPSKSKTDNSLSS-----PLNPKL--WCHVHLKKSLSGSPKVKNSKNSKSPEE 364
QY 334 KLOQ-----XRSSHIQKKAHVALYQLPXGXP--KSTKPL 366
Db 365 HLEEMMKMSPNKLHTNPHIPKG-----PPAKKPGKHSKPL 402
```

Search completed: June 7, 2005, 12:04:51  
Job time : 19.3155 secs

Result No.	Score		Query Match	Length	DB	ID	Description
	Score	Match					
1	4596	100.0	4596	6	A66650	A66650 Sequence 6	
2	4578.4	99.6	4596	6	AX429240	AX429240 Sequence	
3	4578.4	99.6	4596	6	XY774978	XY774978 Sequence	
4	4578.4	99.6	4596	9	HSNIX	Y10256 H.sapiens m	
5	4438.4	96.6	4584	6	CQ723104	CQ723104 Sequence	
6	4249.4	92.5	4452	9	BC035576	BC035576 Homo sapi	
7	3149.6	68.5	3156	6	AR062230	AR062230 Sequence	
8	3149.6	68.5	3156	6	AR063065	AR063065 Sequence	
9	3149.6	68.5	3156	6	AR068515	AR068515 Sequence	
10	3149.6	68.5	3156	6	BD062401	BD062401 NIK prote	
11	2837.6	61.7	2844	6	AR429673	AR429673 Sequence	
12	2873.4	58.2	3152	6	CQ845924	CQ845924 Sequence	
13	2573.4	58.2	3152	9	AK131438	AK131438 Homo sapi	
14	2054.6	44.7	2829	10	AF143094	AF143094 Mus musc	
15	1851.2	40.3	2631	6	A66647	A66647 Sequence 3	
16	1556.2	33.9	48032	9	AC003963	AC003963 Homo sapi	
17	1556.2	33.9	96875	9	AC008105	AC008105 Homo sapi	
18	1556.2	33.9	100836	2	AC087298	AC087298 Homo sapi	
19	1544.2	33.6	142326	2	AC024047	AC024047 Homo sapi	

Db 181 AGTGACCTGCTCTGTGTTGGTCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTG 240  
Qy 241 ATGGAAATGGCTGCTGCCAGAGTCCCTCGCTCAGCAGTGGGCGACGACGAAGAACTCCCC 300  
Db 241 ATGGAAATGGCTGCTGCCAGAGTCCCTCGCTCAGCAGTGGGCGACGACGAAGAACTCCCC 300  
Qy 301 AAGCCAAAGGAGAAAGACGCGCCACTCTGGGGAAGAAACAGAGCTCCGTTCTCAAGCTTCAG 360  
Db 301 AAGCCAAAGGAGAAAGACGCGCCACTCTGGGGAAGAAACAGAGCTCCGTTCTCAAGCTTCAG 360  
Qy 361 GCCGTGGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAAATGAAGTGAATACC 420  
Db 361 GCCGTGGAGAAAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAAATGAAGTGAATACC 420  
Qy 421 AAGGGCAGACCAAGGAAGGCTCCGAGGAGGGCCAGCTGCCATCTCTATCATCGCCAG 480  
Db 421 AAGGGCAGACCAAGGAAGGCTCCGAGGAGGGCCAGCTGCCATCTCTATCATCGCCAG 480  
Qy 481 GCTGAGTGTGAGAAATAGCCAAAGTTTCAGCCCACTTTTCAGAACGCAATTTTCATCGCT 540  
Db 481 GCTGAGTGTGAGAAATAGCCAAAGTTTCAGCCCACTTTTCAGAACGCAATTTTCATCGCT 540  
Qy 541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATTCAGATCCCAACAAATGTGCCCAT 600  
Db 541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATTCAGATCCCAACAAATGTGCCCAT 600  
Qy 601 GCTACAGAGGCAAAATGSCCGTGTGTTGGAGGGAAGCGTCGACGAAGACCCGG 660  
Db 601 GCTACAGAGGCAAAATGSCCGTGTGTTGGAGGGAAGCGTCGACGAAGACCCGG 660  
Qy 661 AAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGGAGTGGCCCTTGGCCAAA 720  
Db 661 AAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGGAGTGGCCCTTGGCCAAA 720  
Qy 721 CCCCTCCCAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGATGAGTCT 780  
Db 721 CCCCTCCCAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGATGAGTCT 780  
Qy 781 CCNCTGGGCGCCCATATGTTAGAAACACCCGAGTTCACCAAGCTCTGAAGGAACCA 840  
Db 781 CCNCTGGGCGCCCATATGTTAGAAACACCCGAGTTCACCAAGCTCTGAAGGAACCA 840  
Qy 841 GGCCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCTACGGCGGCTCTGCTCGA 900  
Db 841 GGCCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCTACGGCGGCTCTGCTCGA 900  
Qy 901 TCAGAACTCCAAACTGATCAGCCCTTGCATATGTGAAACCAAGTGTGAAACTGCAAC 960  
Db 901 TCAGAACTCCAAACTGATCAGCCCTTGCATATGTGAAACCAAGTGTGAAACTGCAAC 960  
Qy 961 CACCCCGAGAGCGAGGCGCCCTGCGCCCTGCCACGACCCCTTCCCTTATAGCAGACTG 1020  
Db 961 CACCCCGAGAGCGAGGCGCCCTGCGCCCTGCCACGACCCCTTCCCTTATAGCAGACTG 1020  
Qy 1021 CCTCATCCCTTCCCAATTCACCCCTCCAGCCCTGGAAACCTCACCCCTTGGAGTCCCTTC 1080  
Db 1021 CCTCATCCCTTCCCAATTCACCCCTCCAGCCCTGGAAACCTCACCCCTTGGAGTCCCTTC 1080  
Qy 1081 CTGGGCAAACTGGCTGTGTAGACAGCGAGAAACCTTGGCTGACCCCACTGAGCAAA 1140  
Db 1081 CTGGGCAAACTGGCTGTGTAGACAGCGAGAAACCTTGGCTGACCCCACTGAGCAAA 1140  
Qy 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCGCCCACTGGAGCCAGCTGCCTG 1200  
Db 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCGCCCACTGGAGCCAGCTGCCTG 1200  
Qy 1201 TCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCTTCGAAGGC 1260  
Db 1201 TCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCTTCGAAGGC 1260  
Qy 1261 AGCGTAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGCGCAAGACCTGGGCGACCG 1320  
Db 1261 AGCGTAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGCGCAAGACCTGGGCGACCG 1320

Qy 1321 GGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGGTGTCTCTGCTCACT 1380  
Db 1321 GGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGGTGTCTCTGCTCACT 1380  
Qy 1381 GAGAAACTCAAGCCAGTGGATTATAGTACCGAGAAAGTCCACTGGGCGACGACACAG 1440  
Db 1381 GAGAAACTCAAGCCAGTGGATTATAGTACCGAGAAAGTCCACTGGGCGACGACACAG 1440  
Qy 1441 CTCGCTGGGCGAGAGCTCTTCGGAGAGTGCACAGGATGAGGACAAAGCACTGGC 1500  
Db 1441 CTCGCTGGGCGAGAGCTCTTCGGAGAGTGCACAGGATGAGGACAAAGCACTGGC 1500  
Qy 1501 TTCCAGTGCCTGTCAAAAAGGTCGCTCGCTGGAAGTATTTTCGGGCGAGGAGCTGATGCA 1560  
Db 1501 TTCCAGTGCCTGTCAAAAAGGTCGCTCGCTGGAAGTATTTTCGGGCGAGGAGCTGATGCA 1560  
Qy 1561 TGTGAGGATTTGACCTCAACCAAGATTTGTCCTTTGTATGGAGCTGTGAGAAAGGCT 1620  
Db 1561 TGTGAGGATTTGACCTCAACCAAGATTTGTCCTTTGTATGGAGCTGTGAGAAAGGCT 1620  
Qy 1621 TGGGTCAACATCTTCATGGAGCTGCTGAAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680  
Db 1621 TGGGTCAACATCTTCATGGAGCTGCTGAAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680  
Qy 1681 CAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGGCTGGAGGGTCTG 1740  
Db 1681 CAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGGCTGGAGGGTCTG 1740  
Qy 1741 GAATACCTCCACTCAGCAAGGATTTCTGATGGGAGCTCAAAAGCTGCAACGTCCTCTG 1800  
Db 1741 GAATACCTCCACTCAGCAAGGATTTCTGATGGGAGCTCAAAAGCTGCAACGTCCTCTG 1800  
Qy 1801 TCCAGCATGGGAGCTGTCAGCTCTGTCACCTTTGGCCATGTGTGTGCTTTCAAGCT 1860  
Db 1801 TCCAGCATGGGAGCTGTCAGCTCTGTCACCTTTGGCCATGTGTGTGCTTTCAAGCT 1860  
Qy 1861 GATGCTCTGGGAAAGTCTCTGCTCAGAGGAGCTACATCTCTGGCACAGAACCCACATG 1920  
Db 1861 GATGCTCTGGGAAAGTCTCTGCTCAGAGGAGCTACATCTCTGGCACAGAACCCACATG 1920  
Qy 1921 GCTCGGAGTGTGCTGGGCGAGAGCTGCAAGGAGGCTGATGCTGAGGAGGCTGTC 1980  
Db 1921 GCTCGGAGTGTGCTGGGCGAGAGCTGCAAGGAGGCTGATGCTGAGGAGGCTGTC 1980  
Qy 1981 TGTATGATGCTGCACATGCTCAAGGCTGCCACCCCTGGAATCTTCTCCGAGGCGG 2040  
Db 1981 TGTATGATGCTGCACATGCTCAAGGCTGCCACCCCTGGAATCTTCTCCGAGGCGG 2040  
Qy 2041 TCTGCTCTCAAGATTTGCGAGAGCTCCGCTGTGAGGAGATCCACCTCTCTGCGCC 2100  
Db 2041 TCTGCTCTCAAGATTTGCGAGAGCTCCGCTGTGAGGAGATCCACCTCTCTGCGCC 2100  
Qy 2101 CCTCTCAGACCCAGGCTCCAGAGGCTGAGGAAAGAGCCCATCCACGCGCTGTCT 2160  
Db 2101 CCTCTCAGACCCAGGCTCCAGAGGCTGAGGAAAGAGCCCATCCACGCGCTGTCT 2160  
Qy 2161 GCAGCGAGCTGGGAGGAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGAGC 2220  
Db 2161 GCAGCGAGCTGGGAGGAGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAGAGC 2220  
Qy 2221 CTTTGGAGGGGAGATATAAAGAACCAAGACATCCACGCGCAATCAAGCCTAATACCAC 2280  
Db 2221 CTTTGGAGGGGAGATATAAAGAACCAAGACATCCACGCGCAATCAAGCCTAATACCAC 2280  
Qy 2281 CAGACCTCTCAGCCAGCGAGAGCTTTGCGCAAGGCGCCAGGCGCCCGGCGAGCT 2340  
Db 2281 CAGACCTCTCAGCCAGCGAGAGCTTTTTCGCAAGGCGCCAGGCGCCCGGCGAGCT 2340  
Qy 2341 GAGGAGACAAAGGAGAGAGCCCTTAAGCTTCAAGCTCTCTCCCAAGAGAGCCCGCAGAG 2400  
Db 2341 GAGGAGACAAAGGAGAGAGCCCTTAAGCTTCAAGCTCTCTCCCAAGAGAGCCCGCAGAG 2400

[illegible][illegible]

[illegible]



Qy	1744	TACCTCCACTCACAAGAGATTCTGCATGGGACGCTCAAAGCTGACAAGCTGCTCTCTGTC	1803
Db	1745	TACCTCCACTCACAAGAGATTCTGCATGGGACGCTCAAAGCTGACAAGCTGCTCTCTGTC	1804
Qy	1804	AGCATGGGAGCCACGACGCCCTCTGTGACATTTGGCCATGCTGTGTCTTCAACCTGAT	1863
Db	1805	AGCATGGGAGCCACGACGCCCTCTGTGACATTTGGCCATGCTGTGTCTTCAACCTGAT	1864
Qy	1864	GGCTTGGGAAAGTCTTGTCTCACAGGAGCTACATCCCTGGCACAGAGACCCACATGGCT	1923
Db	1865	GGCTTGGGAAAGTCTTGTCTCACAGGAGCTACATCCCTGGCACAGAGACCCACATGGCT	1924
Qy	1924	CCGAGGTGTGTCTGGGACAGAGCTGCGAACCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1983
Db	1925	CCGAGGTGTGTCTGGGACAGAGCTGCGAACCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1984
Qy	1984	ATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTGATTTCTTCGAGGCGCGCTC	2043
Db	1985	ATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTGATTTCTTCGAGGCGCGCTC	2044
Qy	2044	TGCTCAAGATTGCCAGGAGCTCCGCTCTGTGAGGAGATCCACACCTCTCGCGCCCT	2103
Db	2045	TGCTCAAGATTGCCAGGAGCTCCGCTCTGTGAGGAGATCCACACCTCTCGCGCCCT	2104
Qy	2104	CTCACAGCCCAAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGTCTGCA	2163
Db	2105	CTCACAGCCCAAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCGTGTCTGCA	2164
Qy	2164	CGGAGCTGGAGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGAGCCT	2223
Db	2165	CGGAGCTGGAGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGAGCCT	2224
Qy	2224	TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACACACAG	2283
Db	2225	TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTACACACAG	2284
Qy	2284	ACCTTCCATGCCAGCGAGAGAGCTTTGCGCAAGAGGCCCCAGGGCCCGCGCCAGCTGAG	2343
Db	2285	ACCTTCCATGCCAGCGAGAGAGCTTTGCGCAAGAGGCCCCAGGGCCCGCGCCAGCTGAG	2344
Qy	2344	GAGACAACAGGACAGAGCCCTTAAGCTCAGCCTCTCTTCCACAGAGCCCGCCAGAGCCA	2403
Db	2345	GAGACAACAGGACAGAGCCCTTAAGCTCAGCCTCTCTTCCACAGAGCCCGCCAGAGCCA	2404
Qy	2404	AACAAGTCTCCTCCCTTGACTTTGAGCAAGAGAGGTCTGGGATGTGGAAACCCCTTACCT	2463
Db	2405	AACAAGTCTCCTCCCTTGACTTTGAGCAAGAGAGGTCTGGGATGTGGAAACCCCTTACCT	2464
Qy	2464	CTGTCTCCTCCCTGGAGCCAGCCCTCGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465	CTGTCTCCTCCCTGGAGCCAGCCCTCGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524	GTCCCGGACAGGAATGTCAGCAGCTGGAATATGAAATATTCCTCAACAGCCTGTCCAG	2583
Db	2525	GTCCCGGACAGGAATGTCAGCAGCTGGAATATGAAATATTCCTCAACAGCCTGTCCAG	2584
Qy	2584	CCATTTTCTCTGGAGGACAGGACAAATTTCTCTGTGCTCAGCATTCGACAGCCTCTCC	2643
Db	2585	CCATTTTCTCTGGAGGACAGGACAAATTTCTCTGTGCTCAGCATTCGACAGCCTCTCC	2644
Qy	2644	CTGTCCGATGACAGTGAGAGAAACCCATCAAGGCCCTCTCAAAGCTCGCGGACACCCCTG	2703
Db	2645	CTGTCCGATGACAGTGAGAGAAACCCATCAAGGCCCTCTCAAAGCTCGCGGACACCCCTG	2704
Qy	2704	AGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTCCAGCTGGAACATG	2763
Db	2705	AGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGAGGCTCGAAGCTCCAGCTGGAACATG	2764
Qy	2764	GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTTATTTCAATGGTGTGAAAGTCAA	2823
Db	2765	GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTTATTTCAATGGTGTGAAAGTCAA	2824

Qy	2824	ATACAGTCTCTTAATGTGGAACACCTGCACATCCGGAGTTTCCACCGGGTCAAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGTGGAACACCTGCACATCCGGAGTTTCCACCGGGTCAAAAGTGGGA	2884
Qy	2884	GACATCCGCCACTGGCATCAGCAGCCAGATCCCAAGCTGCAGCCTTACGTTGGTCAACAAA	2943
Db	2885	GACATCCGCCACTGGCATCAGCAGCCAGATCCCAAGCTGCAGCCTTACGTTGGTCAACAAA	2944
Qy	2944	GACGGGACGCTGTTCCGTACGACATGGAGGTGCAGATCGGGCATCGACTGCAGTGC	3003
Db	2945	GACGGGACGCTGTTCCGTACGACATGGAGGTGCAGATCGGGCATCGACTGCAGTGC	3004
Qy	3004	ACACTGCGCCCTGATGAGCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCAGCTGAG	3063
Db	3005	ACACTGCGCCCTGATGAGCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCAGCTGAG	3064
Qy	3064	AACAGGCCCTAAACCTGCCCTCCACCGCGCTCCACATGCGCGAAGCAGCCTTCCCTG	3123
Db	3065	AACAGGCCCTAAACCTGCCCTCCACCGCGCTCCACATGCGCGAAGCAGCCTTCCCTG	3124
Qy	3124	CTCGGTGCACGATGCTGCCCTGAAAAACACAGGGCTCAGCCGTTCCAGGGGATTGCCAGCC	3183
Db	3125	CTCGGTGCACGATGCTGCCCTGAAAAACACAGGGCTCAGCCGTTCCAGGGGATTGCCAGCC	3184
Qy	3184	CCCCGGTCAAGTGGGAAACAGGGCTTCGACAGCAAGTGGGGGCAAGCAAAATGCC	3243
Db	3185	CCCCGGTCAAGTGGGAAACAGGGCTTCGACAGCAAGTGGGGGCAAGCAAAATGCC	3244
Qy	3244	TCCAGGATTTTCAACCTGAGCCTGCCACCCCTGCTGAAAAAACAATCCCGCCAGTGAA	3303
Db	3245	TCCAGGATTTTCAACCTGAGCCTGCCACCCCTGCTGAAAAAACAATCCCGCCAGTGAA	3304
Qy	3304	GAGACAGAAGGAGATGGCAGGAGTTACTCTGGGAAAAACAAAAAGGGATCTTTTCTGCC	3363
Db	3305	GAGACAGAAGGAGATGGCAGGAGTTACTCTGGGAAAAACAAAAAGGGATCTTTTCTGCC	3364
Qy	3364	CCTGCTCCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGAACATTTGTTGGCAGACAGG	3423
Db	3365	CCTGCTCCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGAACATTTGTTGGCAGACAGG	3424
Qy	3424	GGAGAGCAGCTTCAGCCTGGGTGAGAAAGGGTGGGCGAGCCCTTCGGCCCTCACCCCTC	3483
Db	3425	GGAGAGCAGCTTCAGCCTGGGTGAGAAAGGGTGGGCGAGCCCTTCGGCCCTCACCCCTC	3484
Qy	3484	CAGCTCTCTGTGAGAGTGTCAAGTGTCTAAGGGGCCAAACTCAGGTTTCAAGTGAACCA	3543
Db	3485	CAGCTCTCTGTGAGAGTGTCAAGTGTCTAAGGGGCCAAACTCAGGTTTCAAGTGAACCA	3544
Qy	3544	GGTCAGCAGGTATGCCCGCCCTGAGTTAAGGGGGCCCTCTAAACCCCTTGGCCTGGCCTC	3603
Db	3545	GGTCAGCAGGTATGCCCGCCCTGAGTTAAGGGGGCCCTCTAAACCCCTTGGCCTGGCCTC	3604
Qy	3604	ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATTCCTGACCCCTGGGAAGGCT	3663
Db	3605	ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATTCCTGACCCCTGGGAAGGCT	3664
Qy	3664	CCCTGGTAGAATACACCACTTTTTCAGGTTGTTGCAACACAGCTCTCAGTTCACCTCT	3723
Db	3665	CCCTGGTAGAATACACCACTTTTTCAGGTTGTTGCAACACAGCTCTCAGTTCACCTCT	3724
Qy	3724	GGTTTACGCAAGGACCAAGAAAGGTGTGAAGTGAAGTGTCTTTCAGTCCCGCAGACATGT	3783
Db	3725	GGTTTACGCAAGGACCAAGAAAGGTGTGAAGTGAAGTGTCTTTCAGTCCCGCAGACATGT	3784
Qy	3784	GGCCCTTTGTGTGTGCTACCACTTTTCCAGAGCAGCAGGCCCCCAGGCCCCCTTCAGGC	3843
Db	3785	GGCCCTTTGTGTGTGCTACCACTTTTCCAGAGCAGCAGGCCCCCAGGCCCCCTTCAGGC	3844
Qy	3844	CCAGCACTGCCCCAGACTCGCTGCACTGATGTTCCCTCATCTGTAAGGTTGAAGGTTGAT	3903
Db	3845	CCAGCACTGCCCCAGACTCGCTGCACTGATGTTCCCTCATCTGTAAGGTTGAAGGTTGAT	3904
Qy	3904	GCAGGATATGCTGCACAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAG	3963

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
QY	5 GGGGGACTCTGCGCTGTGGAACTGTAGCTGTGTGA-AGGTGGACTCTGTGTACCTTACCTTTGAGG 63
Db	5 GGGGGACTCTGCGCTGTGGAACTGTAGCTGTGTGAGAGGTGGACTCTGTATTACCAATTGAGG 64
QY	64 ATGTTTGGAGATGATGATGTGTGGCAGAGGCACACATAAAGCAGCAGAGACCTTTTGCC 123
Db	65 ATGTTTGGAGGATGATGATGTGTGGCAGAGGCACACATAAAGCAGCAGAGACCTTTTGCC 124
QY	124 CTTGCTCTTTCTCCCAACCCCAAGGCTGACCTGTGTGTTCTCCAGGTCTGGGATTCTTAAGT 183
Db	125 CTTGCTCTTTCTCCCAACCCCAAGGCTGACCTGTGTGTTCTCCAGGTCTGGGATTCTTAAGT 184
QY	184 GACCTGCTCTGTGTGTTGTTCTCTCAGAGTGAAGCAAGGCTGGAGATGGCAGTGTATG 243
Db	185 GACCTGCTCTGTGTGTTGTTCTCTCAGAGTGAAGCAAGGCTGGAGATGGCAGTGTATG 244
QY	244 GAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGCAGCAGAACTCCCAAG 303
Db	245 GAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGCGCAGCAGAACTCCCAAG 304
QY	304 CCAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCCGTCTCAAGCTTTGAGGCC 363
Db	305 CCAAGGAGAGACGCGCCACTGGGGAAGAAACAGAGCTCCGTCTCAAGCTTTGAGGCC 364
QY	364 GTGAGAAAGAGCCTCTGTGTTCTGCGAAAGTGGGAGATCTGAAATGACGTGATACCAAG 423
Db	365 GTGAGAAAGAGCCTCTGTGTTCTGCGAAAGTGGGAGATCTGAAATGACGTGATACCAAG 424
QY	424 GGCACAGCAGGAAGGCTCCGAGGCGAGGCCAGCTCTCTATCATCGCCAGGCT 483
Db	425 GGCACAGCAGGAAGGCTCCGAGGCGAGGCCAGCTCTCTATCATCGCCAGGCT 484
QY	484 GAGTGTGAGAAATAGCAGAGTTTACGCCACCTTTTTCAGAAACGCAATTTTCATCGCTGG 543
Db	485 GAGTGTGAGAAATAGCAGAGTTTACGCCACCTTTTTCAGAAACGCAATTTTCATCGCTGG 544
QY	544 TCCAAAACAGTACAGCAGTCCGAGAGTCTTGATTCAGATCCCAACAAATGTGGCCCATGCT 603
Db	545 TCCAAAACAGTACAGCAGTCCGAGAGTCTTGATTCAGATCCCAACAAATGTGGCCCATGCT 604
QY	604 ACAGAGGCGAAAATGGCCCGTGTGTGTGGAGGGAAGCGTCGCAGCAAAAGCCCGGAAG 663
Db	605 ACAGAGGCGAAAATGGCCCGTGTGTGTGGAGGGAAGCGTCGCAGCAAAAGCCCGGAAG 664
QY	664 AAACGGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCAGGAGTGGCTTTGGCCAAACCC 723
Db	665 AAACGGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCAGGAGTGGCTTTGGCCAAACCC 724
QY	724 CTTCCCGAGGACCTTGAGCAGGAGAGTGCACCATCCAGTGCAGAGAGTGAATCTTCCA 783
Db	725 CTTCCCGAGGACCTTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGAGTGAATCTTCCA 784
QY	784 CTCGGCGCCCATATGTTTGAAGAACACCCCGAGTTTACCAAGGCTCTGAAAGGAACACAGGC 843
Db	785 CTCGGCGCCCATATGTTTGAAGAACACCCCGAGTTTACCAAGGCTCTGAAAGGAACACAGGC 844
QY	844 CTTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGGCCGGCTCTGCTTCGATCA 903
Db	845 CTTTGGGCAACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGGCCGGCTCTGCTTCGATCA 904
QY	904 GAACTCCCAAACTGATCAGCCCCCTTGCATATGTGAAACCACTGTGGAAACTGCAACAC 963
Db	905 GAACTCCCAAACTGATCAGCCCCCTTGCATATGTGAAACCACTGTGGAAACTGCAACAC 964
QY	964 CCCCAGGAGGAGGCGCCCTGCGCCCTGCCACGACACCCCTTCCCTATAGCAGCTGCGCT 1023
Db	965 CCCCAGGAGGAGGCGCCCTGCGCCCTGCCACGACACCCCTTCCCTATAGCAGCTGCGCT 1024
QY	1024 CATCCCTTCCCATTCACCCCTCTCCAGCCCTGGAACCTCACCTCTGGAGTCTTCTGCT 1083
Db	1025 CATCCCTTCCCATTCACCCCTCTCCAGCCCTGGAACCTCACCTCTGGAGTCTTCTGCT 1084

Query Match	99.6%	Score 4578.4	DB 6	Length 4596
Best Local Similarity	100.0%	Pred. No. 0		

Qy	1084	GGCAAACTGGCCCTGTGTAGACGACGCAAAAACCCCTTGCTGTGACCCACACCTGAGCAAACTG	1143
Db	1085	GGCAAACTGGCCCTGTGTAGACGACGCAAAAACCCCTTGCTGTGACCCACACCTGAGCAAACTG	1144
Qy	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCGAGCTGCTGTCT	1203
Db	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCGAGCTGCTGTCT	1204
Qy	1204	CGTGGTCCCATGAGAACTTTTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGCAGC	1263
Db	1205	CGTGGTCCCATGAGAACTTTTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGCAGC	1264
Qy	1264	GTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGCCAAAGACTTGGGAGAGCAGGGGC	1323
Db	1265	GTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGCCAAAGACTTGGGAGAGCAGGGGC	1324
Qy	1324	TCCAGATCCCGGGAGCCAGGCCCAAAACTCAGGACACAGAGGTGCTCTCACTGAG	1383
Db	1325	TCCAGATCCCGGGAGCCAGGCCCAAAACTCAGGACACAGAGGTGCTCTCACTGAG	1384
Qy	1384	AAACTCAAGCCAGTGGATTATAGGTACCGAGAAAGTCCACTGGGCCACGACACAGCTC	1443
Db	1385	AAACTCAAGCCAGTGGATTATAGGTACCGAGAAAGTCCACTGGGCCACGACACAGCTC	1444
Qy	1444	CGCTGGGCAAGGCTCCTTCGAGAGGTGCACAGGATGGAGGACAGAGACTGGCTTC	1503
Db	1445	CGCTGGGCAAGGCTCCTTCGAGAGGTGCACAGGATGGAGGACAGAGACTGGCTTC	1504
Qy	1504	CAGTGGCTGTCAAAAAGGTGGCTGGAAATTTTCGGGCAGAGGAGCTGATGGCATGT	1563
Db	1505	CAGTGGCTGTCAAAAAGGTGGCTGGAAATTTTCGGGCAGAGGAGCTGATGGCATGT	1564
Qy	1564	GCAGGATTGACCTCACCCAGAAATTTCTCTTTGATGGAGCTGTGAGAAAGGCTTGG	1623
Db	1565	GCAGGATTGACCTCACCCAGAAATTTCTCTTTGATGGAGCTGTGAGAAAGGCTTGG	1624
Qy	1624	GTCAACATCTTTCATGGAGCTGCTGGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG	1683
Db	1625	GTCAACATCTTTCATGGAGCTGCTGGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG	1684
Qy	1684	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCTGGAGGCTTGAA	1743
Db	1685	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCTGGAGGCTTGAA	1744
Qy	1744	TACCTCACTCAGAAAGGATTTCTGATGGGACGCTCAAAGCTGACAACTGCTCTCTGTC	1803
Db	1745	TACCTCACTCAGAAAGGATTTCTGATGGGACGCTCAAAGCTGACAACTGCTCTCTGTC	1804
Qy	1804	AGCATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGAT	1863
Db	1805	AGCATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGAT	1864
Qy	1864	GGCTTGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTGGCCACAGACCCATGGCT	1923
Db	1865	GGCTTGGAAAGTCTTGTCTCAGAGGAGCTACATCCCTGGCCACAGACCCATGGCT	1924
Qy	1924	CCGAGGTGGTGTCTGGGACAGGAGTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1983
Db	1925	CCGAGGTGGTGTCTGGGACAGGAGTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1984
Qy	1984	ATGATGTGACATGCTCAACGGCTGCCACCCCTGGACTCAGTTTCTTCGAGGGCCGCTC	2043
Db	1985	ATGATGTGACATGCTCAACGGCTGCCACCCCTGGACTCAGTTTCTTCGAGGGCCGCTC	2044
Qy	2044	TGCCTCAAGATTGCAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTCTGGCCCCCT	2103
Db	2045	TGCCTCAAGATTGCAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTCTGGCCCCCT	2104
Qy	2104	CTCACGCCAGGCCATCCAAAGAGGGGTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2163
Db	2105	CTCACGCCAGGCCATCCAAAGAGGGGTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2164

Qy	2164	GCGAGCTGGAGGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGCCCT	2223
Db	2165	GCGAGCTGGAGGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGCCCT	2224
Qy	2224	TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAATCAAGCAAATTTACCACCAG	2283
Db	2225	TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAATCAAGCCAAATTTACCACCAG	2284
Qy	2284	ACCTTCCATGCCAGCCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGGCCAGCTGAG	2343
Db	2285	ACCTTCCATGCCAGCCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGGCCAGCTGAG	2344
Qy	2344	GAGACAAGCCAGAGAGCCCTTAAGCTCAGGCTCTCTCCACACAGAGCCCCAGAGGCA	2403
Db	2345	GAGACAAGCCAGAGAGCCCTTAAGCTCAGGCTCTCTCCACACAGAGCCCCAGAGGCA	2404
Qy	2404	AACAAGTCTCTCTCTTTCGCTTTGAGCAAGGAGAGTCTGGGATGTGGGAAACCTTTACCT	2463
Db	2405	AACAAGTCTCTCTCTTTCGCTTTGAGCAAGGAGAGTCTGGGATGTGGGAAACCTTTACCT	2464
Qy	2464	CTGTCTCTCTTGGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465	CTGTCTCTCTTGGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524	GTCCCGAGAGGAACTGACAGCTGGAATAAGAAATTTCTCTCAACAGGCTGTCTCCAG	2583
Db	2525	GTCCCGAGAGGAACTGACAGCTGGAATAAGAAATTTCTCTCAACAGGCTGTCTCCAG	2584
Qy	2584	CCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTCAGCATCGAGGCTCTCC	2643
Db	2585	CCATTTTCTCTGGAGGAGCAGGAGCAATTTCTCTGCTCAGCATCGAGGCTCTCC	2644
Qy	2644	CTGTCCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG	2703
Db	2645	CTGTCCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG	2704
Qy	2704	AGCTCAGGCTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCAGCTGGAACATG	2763
Db	2705	AGCTCAGGCTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCAGCTGGAACATG	2764
Qy	2764	GTCTGGCCCGGGGGCCCGACACCCAGCTATTTCATGCTGTAAGTCCAA	2823
Db	2765	GTCTGGCCCGGGGGCCCGACACCCAGCTATTTCATGCTGTAAGTCCAA	2824
Qy	2824	ATACAGTCTCTTAATGTGAACACCTGCACATCCGGAGTTCCACCGGTCGAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGTGAACACCTGCACATCCGGAGTTCCACCGGTCGAAAGTGGGA	2884
Qy	2884	GACATGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACAA	2943
Db	2885	GACATGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACAA	2944
Qy	2944	GACGGGAGCTTGTTCGTACGATGAGGTGCCAGACTCGGGCATCGACTGAGTGC	3003
Db	2945	GACGGGAGCTTGTTCGTACGATGAGGTGCCAGACTCGGGCATCGACTGAGTGC	3004
Qy	3004	ACACTGGCCCTGATGTCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCAGCTGGAG	3063
Db	3005	ACACTGGCCCTGATGTCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCAGCTGGAG	3064
Qy	3064	AACAGGCCCTTAACCTTCCCTTCCACCGGCTCCACATGCTCGGAAAGCAGCTTCTG	3123
Db	3065	AACAGGCCCTTAACCTTCCCTTCCACCGGCTCCACATGCTCGGAAAGCAGCTTCTG	3124
Qy	3124	CTGGTGCAGATGCTGCTTGAATAACAGAGCTCAGCTTCCAGGGATTTGCCAGCC	3183
Db	3125	CTGGTGCAGATGCTGCTTGAATAACAGAGCTCAGCTTCCAGGGATTTGCCAGCC	3184
Qy	3184	CCCGGCTCAGTGGGAAACAGGGCTTCGAGCAGCAAGGTGGGGCAAGCAAGTGGCC	3243
Db	3185	CCCGGCTCAGTGGGAAACAGGGCTTCGAGCAGCAAGGTGGGGCAAGCAAGTGGCC	3244
Qy	3244	TCCAGGATTTCACTGAGCCCTGCCCAACCTGCTGTAATAAATCCGCCACGTGAA	3303

Db 3245 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTCTGCTGAATAAACATCCGCCAGGTGA 3304  
QY 3304 GAGACAGAGGAGGATGCGAGAGTTACCTGGGAAACAAAACAGGAGTCTTTTCTGCC 3363  
Db 3305 GAGACAGAGGAGGATGCGAGAGTTACCTGGGAAACAAAACAGGAGTCTTTTCTGCC 3364  
QY 3364 CTTGCTCAGATCGAGTTGGCTTGACCCGTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423  
Db 3365 CTTGCTCAGATCGAGTTGGCTTGACCCGTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424  
QY 3424 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCTC 3483  
Db 3425 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCCCTC 3484  
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCCAAACTCAGGTTCAAGTGCAGAACCA 3543  
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCCAAACTCAGGTTCAAGTGCAGAACCA 3544  
QY 3544 GGTACAGAGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTTGGCTGCCCTC 3603  
Db 3545 GGTACAGAGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTTGGCTGCCCTC 3604  
QY 3604 ACCTGCGCAGCTCACCCCTTTTGGGTAGGGGAAAAGAAATGCCCTGACCCCTGGGAAGGCT 3663  
Db 3605 ACCTGCGCAGCTCACCCCTTTTGGGTAGGGGAAAAGAAATGCCCTGACCCCTGGGAAGGCT 3664  
QY 3664 CCCTGGTAGAATPACCAACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3723  
Db 3665 CCCTGGTAGAATPACCAACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3724  
QY 3724 GGTTCAGCAAGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3783  
Db 3725 GGTTCAGCAAGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3784  
QY 3784 GCCCTTTCCTGCTGCTACCACTCTTCCAGAGCAGCAGGCCCGAGCCCTTCAGGC 3843  
Db 3785 GCCCTTTCCTGCTGCTACCACTCTTCCAGAGCAGCAGGCCCGAGCCCTTCAGGC 3844  
QY 3844 CCAGCACTCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGAT 3903  
Db 3845 CCAGCACTCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGAT 3904  
QY 3904 GAGGATATGCTCAGAGAACAGTCTGGATGGATGATCATCATGATGCTAAGGAAGCAG 3963  
Db 3905 GCAGGATATGCTCAGAGAACAGTCTGGATGGATGATCATCATGATGCTAAGGAAGCAG 3964  
QY 3964 CAGAGAGAGAGTCCGGCGCCCGAGCCCACTATCAGTGTCCAGCGTGTGGTTCCCCAG 4023  
Db 3965 CAGAGAGAGAGTCCGGCGCCCGAGCCCACTATCAGTGTCCAGCGTGTGGTTCCCCAG 4024  
QY 4024 AGCAGAGCTCAGCATCACACTGACATCACTACCTGCTGCCCTGGCCGAGAGGTACTGC 4083  
Db 4025 AGCAGAGCTCAGCATCACACTGACATCACTACCTGCTGCCCTGGCCGAGAGGTACTGC 4084  
QY 4084 CGAGCGCACTTTGCATCTGTAGTACCTCAAGACATTTTCATGCTCCCTCTGGCAGGCG 4143  
Db 4085 CGAGCGCACTTTGCATCTGTAGTACCTCAAGACATTTTCATGCTCCCTCTGGCAGGCG 4144  
QY 4144 AGGGCAGGCGAGTGACATGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA 4203  
Db 4145 AGGGCAGGCGAGTGACATGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA 4204  
QY 4204 GTCTTGAGCTGTCCATGATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGAAAT 4263  
Db 4205 GTCTTGAGCTGTCCATGATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGAAAT 4264  
QY 4264 GCACCCCTTCCCATTTGCCAGCTTAGCCTCTTCCAGGGGAGCTACTCAGGACTC 4323  
Db 4265 GCACCCCTTCCCATTTGCCAGCTTAGCCTCTTCCAGGGGAGCTACTCAGGACTC 4324  
QY 4324 ACCTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGTAGCCTCAACCTCTCTGGG 4383

Db 4325 ACGTAGCATTAATCAGCTGTGAATCTGTGAGGGGTGTCTGTAGCCTCAACCTCTGGG 4384  
QY 4384 GCAGGGACGCCGAGAGCTCCGTGGGAGAGCTATCCCCACATCTTCCCAAGACAGGCTT 4443  
Db 4385 GCAGGGACGCCGAGAGCTCCGTGGGAGAGCTATCCCCACATCTTCCCAAGACAGGCTT 4444  
QY 4444 TGTCCAGCTGTCCACATTTGAGTCTGCTCCCGGGGAGAGAGCCCGGCCGCCAGCAC 4503  
Db 4445 TGTCCAGCTGTCCACATTTGAGTCTGCTCCCGGGGAGAGAGCCCGGCCGCCAGCAC 4504  
QY 4504 ATAAAGAACTGCAGCTTGGTGTGCTGAGAGTCTGGGTGTGAGAGAACTCTTTGTAAGCAA 4563  
Db 4505 ATAAAGAACTGCAGCTTGGTGTGCTGAGAGTCTGGGTGTGAGAGAACTCTTTGTAAGCAA 4564  
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595  
Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 4

HSNIK HSNIK 4596 bp mRNA linear PRI 13-JAN-1998

LOCUS H.sapiens mRNA for serine/threonine protein kinase, NIK.

DEFINITION Y10256

ACCESSION Y10256.1 GI:1841433

VERSION MAP kinase; NIK protein; serine/threonine protein kinase.

KEYWORDS MAP kinase; NIK protein; serine/threonine protein kinase.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Malinin,N.L., Boldin,M.P., Kovalenko,A.V. and Wallach,D.

TITLE MAP3K-related kinase involved in NF-kappaB induction by TNF, CD95 and IL-1

JOURNAL Nature 385 (6616), 540-544 (1997)

MEDLINE 97172277

PUBMED 9020361

REFERENCE 2 (bases 1 to 4596)

AUTHORS Wallach,D.

TITLE Direct Submission

JOURNAL Submitted (23-DEC-1996) D. Wallach, The Weizmann Institute, Dept of Membrane Research & Biophysics, Rehovot 76100, ISRAEL

COMMENT NIK is a serine/threonine protein-kinase, resembling several MAP kinase kinase kinases (MAP3K), that binds specifically to TRAF2, an adapter proteins associated, either directly or through interaction with other adapter proteins, with several receptors of the TNF/NGF family. NIK overexpression in cells activates the transcription factor NF kappa B. Cellular expression of kinase-deficient NIK-mutants blocks NF kappa B induction by TNF, by either of the two TNF receptors, by CD95 (Fas/Apo-1) and by TRADD, RIP and MORT1/FADD, adapter proteins that bind to these receptors. It also blocks NF kappa B induction by IL-1.

FEATURES

location/Qualifiers

1..4596 /organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/cell\_type="B cells"

misc\_feature 137..149 /note="lysine-rich region"

233..3076 /codon\_start=1

/product="NIK, serine/threonine protein-kinase"

/protein\_id="CAA71306.1"

/db\_xref="GI:1841434"

/db\_xref="GOA:Q99558"

/db\_xref="UniProt/Swiss-Prot:Q99558"

/translation="MAVMEACPGAPGSAVGQKSLPKPKETPKLKKQSVYKLEA VEKSPVFCGKEILNDVITKGTAGESEAGAPAAISIIAQACENQSEFPTSERIFI AGSKQYSSQESLDQIPNNVAHATEGKMARVCKWKKRRKARRKKKSKSLAHAGVA LAKPLPRPEQESCTIPVQDESPGAPVVRNTPQTPLEKPLKEPGLQCLFQKLGELR PALPRLSELHLKLSIPLOCLNHWKLHPDGGPLPLPTHPFPYSRLPHPPPLQPLQPK PHLPSFLGKLUACVDSQKPLPDPHLUSKLACVDSKPLPLGPHLEPSCLSRGAHEKFSVE

EYLVAHQGVSSQAHSLTSLAKTWAARGSRSPKPTEDNEGVLLTEKLKPDVYE  
YREYVWATHQRLGRGSEVHEMRDQFQCAVKVRLEVPRAELMACAGLTSP  
RIVLYGAVREGPWNTFMELLEGSLQOLVKEQCCLPEDRALYLLQOALEGLYLSR  
RRIILGDVADNVLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDIIPGTEHMAPE  
VVLGRSDAKVDVSSCMMLHLNGCHPWTOFFRGPLCLKIASERPVPREIPSPCAP  
LTAQAIQEGLEKEPIHVRVSAELGKVNRLAQVGLKSPWRGKYKEPRHPPPNOAY  
EPTLHAQPRELSPRAPGRPAEBETVGRAPKLOPLPEPPENKSPPLTUSKEESGMW  
EPILSLSDSEKNPSKASQSRDTLSSVHSSWQAEARSSMMVLAERPRDPTDS  
YFNGVKVQIOISLNGEHLHIREPHRVKVGDIATGISSQIPAAAPSLVTKDGPVRYDME  
VPDSGIDLQCTLAPDGSFANRKYKHQLENRP"

misc\_feature

/note=".653

misc\_feature

/note="kinase domain"

711..729

/note="proline-rich region"

ORIGIN

Query Match 99.6%; Score 4578.4; DB 9; Length 4596;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	5	GGGGGACTGTCCCGTGTGGAACGTGTAGCTGTGGA-AGGTGGACTCTGTTTACCATTGAGG	63
Db	5	GGGGGACTGTCCCGTGTGGAACGTGTAGCTGTGAGAGGTGGACTCTGTTTACCATTGAGG	64
Qy	64	ATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGCAGACCCCTTTGGCC	123
Db	65	ATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGCAGACCCCTTTGGCC	124
Qy	124	CTCGCTTTTCCCCCAACCCAGGCTGACCTGTGTTCTCCAGAGTCTGGGATCTAAGT	183
Db	125	CTCGCTTTTCCCCCAACCCAGGCTGACCTGTGTTCTCCAGAGTCTGGGATCTAAGT	184
Qy	184	GACCTGCTCTGTGTTTGTGCTCTCTCAGGATGAGCAAGCTCGGAGATGGCAGTGATG	243
Db	185	GACCTGCTCTGTGTTTGTGCTCTCTCAGGATGAGCAAGCTCGGAGATGGCAGTGATG	244
Qy	244	GAATATGGCTTCCCGAGTGCCCTGGCTCAGCAGTGGGCGACAGAAAGAACTCCCCCAAG	303
Db	245	GAATATGGCTTCCCGAGTGCCCTGGCTCAGCAGTGGGCGACAGAAAGAACTCCCCCAAG	304
Qy	304	CCAAAGGAGAAAGACCGCCGCACTGGGGAAGAAAAGAGCTCCGCTCAAGCTTGAGGCC	363
Db	305	CCAAAGGAGAAAGACCGCCGCACTGGGGAAGAAAAGAGCTCCGCTCAAGCTTGAGGCC	364
Qy	364	GTGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCCTGAATGACGTGATTAACCAAG	423
Db	365	GTGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCCTGAATGACGTGATTAACCAAG	424
Qy	424	GGCACAGCCAAAGGAGGCTCCGAGGCGAGGGCCAGCTGCCATCTCTATCATCGCCCAAGCT	483
Db	425	GGCACAGCCAAAGGAGGCTCCGAGGCGAGGGCCAGCTGCCATCTCTATCATCGCCCAAGCT	484
Qy	484	GAGTGTGAGAAATAGCAAGAGTTGAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG	543
Db	485	GAGTGTGAGAAATAGCAAGAGTTGAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG	544
Qy	544	TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACATGTGGCCCATGCT	603
Db	545	TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACATGTGGCCCATGCT	604
Qy	604	ACAGAGGCAAAATGGCCCGTGTGTGTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	663
Db	605	ACAGAGGCAAAATGGCCCGTGTGTGTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	664
Qy	664	AAACGGAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCTTGGCCCAACCC	723
Db	665	AAACGGAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCTTGGCCCAACCC	724
Qy	724	CTCCCCAGGACCCCTGAGCAGGAGAGTGCACCAATCCCAAGTGCAGGAGGATGAGTCTCCA	783
Db	725	CTCCCCAGGACCCCTGAGCAGGAGAGTGCACCAATCCCAAGTGCAGGAGGATGAGTCTCCA	784

1865 GGCCTGGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCT 1924  
1924 CCGGAGGTGGTGTCTGGCAGAGCTTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGGTGT 1983  
1925 CCGGAGGTGGTGTCTGGCAGAGCTTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGGTGT 1984  
1984 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCCTTCGAGGGCCGCTC 2043  
1985 ATGATGCTGCACATGCTCAA CGGCTGCCACCCCTGGACTCAGTTCCTTCGAGGGCCGCTC 2044  
2044 TGCCCTCAAGATTGCCAGCAGCCTCCGCTGTGAGGGAGTCCACCCCTCTCTGGCCCT 2103  
2045 TGCCCTCAAGATTGCCAGCAGCCTCCGCTGTGAGGGAGTCCACCCCTCTCTGGCCCT 2104  
2104 CTCACAGCCCGCCCATCCAGAGGGGTGAGGAAGGCCATCCACGGGTGTCTGCA 2163  
2105 CTCACAGCCCGCCCATCCAGAGGGGTGAGGAAGGCCATCCACGGGTGTCTGCA 2164  
2164 GCGGAGCTGGGAGGAAGTGAACCGGCACCTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223  
2165 GCGGAGCTGGGAGGAAGTGAACCGGCACCTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2224  
2224 TGGAGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACACCAG 2283  
2225 TGGAGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACACCAG 2284  
2284 ACCCTCCATGCCAGCGAGAGCTTTGCCAAGGGGCCACAGGCCCGGGCCAGCTGAG 2343  
2285 ACCCTCCATGCCAGCGAGAGCTTTGCCAAGGGGCCACAGGCCCGGGCCAGCTGAG 2344  
2344 GAGACAAAGGACAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCGCCACAGGCA 2403  
2345 GAGACAAAGGACAGAGCCCTTAAGCTCCAGCCTCTCTCTCCACAGAGCCGCCACAGGCA 2404  
2404 AACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT 2463  
2405 AACAAAGTCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT 2464  
2464 CTGTCTCTCTGGAGCAGCCCTTGCAGAAACCCAGTCCACAGAGGGGGAAGCAACC 2523  
2465 CTGTCTCTCTGGAGCAGCCCTTGCAGAAACCCAGTCCACAGAGGGGGAAGCAACC 2524  
2524 GTCCGGAGCAGAACTGCAGAGCTGGAAATAGAAATTATCTCAACAGCCTGTCCAG 2583  
2525 GTCCGGAGCAGAACTGCAGAGCTGGAAATAGAAATTATCTCAACAGCCTGTCCAG 2584  
2584 CCATTTTCTCTGGAGGAGCAGGACAAATTCTCTCGTGCTCAGCATCGACAGCCTCTCC 2643  
2585 CCATTTTCTCTGGAGGAGCAGGACAAATTCTCTCGTGCTCAGCATCGACAGCCTCTCC 2644  
2644 CTGTGGATGACAGTGAGAAAGCCATCAAGGCTCTCAAGCTCGGGGACACCCCTG 2703  
2645 CTGTGGATGACAGTGAGAAAGCCATCAAGGCTCTCAAGCTCGGGGACACCCCTG 2704  
2704 AGCTCAGGCTACATCTCTGGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGAAACATG 2763  
2705 AGCTCAGGCTACATCTCTGGAGCAGCAGGCGGAGCTCGAAGCTCCAGCTGGAAACATG 2764  
2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAA 2823  
2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAA 2824  
2824 ATACAGTCTCTTAATGGTGAACCTGACATCCGGAGTTCACCGGCTCAAGTGGGA 2883  
2825 ATACAGTCTCTTAATGGTGAACCTGACATCCGGAGTTCACCGGCTCAAGTGGGA 2884  
2884 GACATCGCAGCTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGTTGGTTCACCAA 2943  
2885 GACATCGCAGCTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGTTGGTTCACCAA 2944  
2944 GACGGCAGCCTGTTCCGTACGACATGGAGGTGCCAGATTCGGGCATCGACCTGCAAGTGC 3003

Db 2945 GACGGCAGCCTGTTCTGCTACGACATGAGGTGCCAGACTCGGGATCGACCTGCAAGTGC 3004  
Qy 3004 ACATCTGGCCCTGATGTCAGCTTCGCTTGGAGCTTCAAGCTGAGGGTCAAGCTATGGCCAGTGGAG 3063  
Db 3005 ACATCTGGCCCTGATGTCAGCTTCGCTTGGAGCTTCAAGCTATGGCCAGTGGAGTGGAG 3064  
Qy 3064 AACAGGCCCTAAACCTTGCCTTCCACCGCGGCTCCACACTCGCGGAAAGCAGCTTCCTG 3123  
Db 3065 AACAGGCCCTAAACCTTGCCTTCCACCGCGGCTCCACACTCGCGGAAAGCAGCTTCCTG 3124  
Qy 3124 CTGCGTGCACATGCTGCGCTTGAACACACAGGCTTCCAGCGCTTCCAGGGATTCGCAGCC 3183  
Db 3125 CTGCGTGCACATGCTGCGCTTGAACACACAGGCTTCCAGCGCTTCCAGGGATTCGCAGCC 3184  
Qy 3184 CCCCGGCTCAAGTGGGAAACAGGCGCTTCGAGCAGCAAGGTGGGGCAAGCAGAAATGCC 3243  
Db 3185 CCCCGGCTCAAGTGGGAAACAGGCGCTTCGAGCAGCAAGGTGGGGCAAGCAGAAATGCC 3244  
Qy 3244 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTGTGAAAAAACAATCCGCCACGTGAA 3303  
Db 3245 TCCAGGATTTTCAACCTGAGCCCTGCCCCACCTGTGAAAAAACAATCCGCCACGTGAA 3304  
Qy 3304 GAGACAGAGAGATGTCAGAGTTTACCTGGGAAACAAACAGGGATCTTTTCTGCC 3363  
Db 3305 GAGACAGAGAGATGTCAGAGTTTACCTGGGAAACAAACAGGGATCTTTTCTGCC 3364  
Qy 3364 CCTGCTCCAGTTCGAGTTGGCTTGGCTTGGATTCAGTGACCAATTTGTGGCAGACAGG 3423  
Db 3365 CCTGCTCCAGTTCGAGTTGGCTTGGCTTGGATTCAGTGACCAATTTGTGGCAGACAGG 3424  
Qy 3424 GGAGACAGCTTTCAGAGCTGGGTGAGAGGGTGGGGAGCCCTTCGGCCCTTCAACCCTC 3483  
Db 3425 GGAGACAGCTTTCAGAGCTGGGTGAGAGGGTGGGGAGCCCTTCGGCCCTTCAACCCTC 3484  
Qy 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAACTCAGTTTCAAGTGCAGAACCA 3543  
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAACTCAGTTTCAAGTGCAGAACCA 3544  
Qy 3544 GGTCAAGAGTATGCCCGCTTGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3603  
Db 3545 GGTCAAGAGTATGCCCGCTTGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3604  
Qy 3604 ACCTGGCAGCTACCCCTTTTGGGTGAGGGGAAAAAGAAATGCTTGAACCTTGGGAAAGGT 3663  
Db 3605 ACCTGGCAGCTACCCCTTTTGGGTGAGGGGAAAAAGAAATGCTTGAACCTTGGGAAAGGT 3664  
Qy 3664 CCCTGGTGAATAACACACATTTTTCAGTTTGTGCAACAGAGTCCCTGAGTTGACCTCT 3723  
Db 3665 CCCTGGTGAATAACACACATTTTTCAGTTTGTGCAACAGAGTCCCTGAGTTGACCTCT 3724  
Qy 3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3783  
Db 3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3784  
Qy 3784 GCGCCCTTGTGCTGGCTTACCACTTTTCCAGAGCAGAGGCCCGGAGCCCTTCAGGC 3843  
Db 3785 GCGCCCTTGTGCTGGCTTACCACTTTTCCAGAGCAGAGGCCCGGAGCCCTTCAGGC 3844  
Qy 3844 CCAGCAGTTCGCCCCAGACTCGCTGGCAGTTCAGTTCCCTCATCTGTAAAGTGAAGGTGAT 3903  
Db 3845 CCAGCAGTTCGCCCCAGACTCGCTGGCAGTTCAGTTCCCTCATCTGTAAAGTGAAGGTGAT 3904  
Qy 3904 CGAGGATATGCTGCAGGAAACAGTCTGTGGATGGACATGATCAGTGTCAAGGAAAGCAG 3963  
Db 3905 CGAGGATATGCTGCAGGAAACAGTCTGTGGATGGACATGATCAGTGTCAAGGAAAGCAG 3964  
Qy 3964 CAGAGAGAGAGTCCGGGGCCCAAGCCCACTATCAGTGTCCAGCGTGTGTTCCCGCAG 4023  
Db 3965 CAGAGAGAGAGTCCGGGGCCCAAGCCCACTATCAGTGTCCAGCGTGTGTTCCCGCAG 4024  
Qy 4024 AGCAGAGCTCAGCATCAGTGCAGACTCACCTTGCCTGCCCTTGGCCAGAGGGTACTGC 4083  
Db 4025 AGCAGAGCTCAGCATCAGTGCAGACTCACCTTGCCTGCCCTTGGCCAGAGGGTACTGC 4084





QY	1202	CTCGTGGTCCCATGAGAAGTTTTCTGTGGAGGAATACCTAGTCATCTCTGCAAGGCA	1261	2282	AGACCTCATGCTCCAGCCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGCGCAGCTG	2341
Db	1203	CTCGTGGTGGCCATGAGAGTTTTCTGTGGAGGAATACCTAGTCATCTCTGCAAGGCA	1262	2283	AGACCTCATGCTCCAGCCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCCGCGCAGCTG	2342
QY	1262	GGTGAAGTCAAGCCAGGCGCCACAGCCTGACAGCCTGGCCCAAGACCTGGGACAGCGG	1321	2342	AGAGAGAAACAGGACAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAGCCCCAGAGC	2401
Db	1263	GGTGAAGTCAAGCCAGGCGCCACAGCCTGACAGCCTGGCCCAAGACCTGGGACAGCGG	1322	2343	AGAGAGAAACAGGACAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAGCCCCAGAGC	2402
QY	1322	GCTCCAGATCCGGGAGCCAGCCCAAACTGAGGACAAAGAGGGTCTCTGCTCACTG	1381	2402	CAAAAGATCTCTCTCTGCTTGAATTTGAGCAAGAGGAGTCTGGATGTGGGAACCTTTAC	2461
Db	1323	GCTCCAGATCCGGGAGCCAGCCCAAACTGAGGACAAAGAGGGTCTCTGCTCACTG	1382	2403	CAAAAGATCTCTCTCTGCTTGAATTTGAGCAAGAGGAGTCTGGATGTGGGAACCTTTAC	2462
QY	1382	AGAACTCAAGCAGTGGATTATAGTACCGAGAAAGAGTCCATGCGGCCACGACAGC	1441	2462	CTCTGCTCTCTCTGAGCCAGCCCTTGCAGAAACCCAGCTCAACAGAGCGGAAAGCAA	2521
Db	1383	AGAACTCAAGCAGTGGATTATAGTACCGAGAAAGAGTCCATGCGGCCACGACAGC	1442	2463	CTCTGCTCTCTCTGAGCCAGCCCTTGCAGAAACCCAGCTCAACAGAGCGGAAAGCAA	2522
QY	1442	TCCGCTGGGACAGGCTCTTTCGAGAGGTGCACAGGATGGAGACAGACACTGGCT	1501	2522	CCGTCGCGAGCAGGAACTGACAGAGCTGGAAATAGAAATTTCTCTCAACAGCTGTGCC	2581
Db	1443	TCCGCTGGGACAGGCTCTTTCGAGAGGTGCACAGGATGGAGACAGACACTGGCT	1502	2523	CCGTCGCGAGCAGGAACTGACAGAGCTGGAAATAGAAATTTCTCTCAACAGCTGTGCC	2582
QY	1502	TCCAGTGCCTCTCAAAAGGTGCGCTGGAAAGTATTTTCGGGCAGAGAGCTGTGGCAT	1561	2582	AGCCATTTTCTCTGGAGAGCAGGACAAATTTCTCTGTCCTCAGCATCGACAGCTCT	2641
Db	1503	TCCAGTGCCTCTCAAAAGGTGCGCTGGAAAGTATTTTCGGGCAGAGAGCTGTGGCAT	1562	2583	AGCCATTTTCTCTGGAGAGCAGGACAAATTTCTCTGTCCTCAGCATCGACAGCTCT	2642
QY	1562	GTCCAGGATGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGGGCCTT	1621	2642	CCCTGTGCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCC	2701
Db	1563	GTCCAGGATGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGGGCCTT	1622	2643	CCCTGTGCGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCC	2702
QY	1622	GGGTCAACATCTTCATGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGTGCAAGGAGC	1681	2702	TGAGCTCAGGGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCCAGCTGGAAACA	2761
Db	1623	GGGTCAACATCTTCATGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGTGCAAGGAGC	1682	2703	TGAGCTCAGGGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCCAGCTGGAAACA	2762
QY	1682	AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTCTGCGGCCAGGCCCTGGAGGGTCTGG	1741	2762	TGAGCTCAGGGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCCAGCTGGAAACA	2821
Db	1683	AGGGCTGTCTCCAGAGGACCGGGCCCTGTACTCTGCGGCCAGGCCCTGGAGGGTCTGG	1742	2763	TGAGCTCAGGGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCCAGCTGGAAACA	2822
QY	1742	AATACCTCCATCACAAGAGTTCTGCATGGGACGTCAAGCTGCAACAGTGTCTCTGT	1801	2822	AAATACAGTCTCTTAATGTGAACACCTCGACATCCCGGGAGTTCACCGGGTCAAGAGTG	2881
Db	1743	AATACCTCCATCACAAGAGTTCTGCATGGGACGTCAAGCTGCAACAGTGTCTCTGT	1802	2823	AAATACAGTCTCTTAATGTGAACACCTCGACATCCCGGGAGTTCACCGGGTCAAGAGTG	2882
QY	1802	CCAGGATGGAGCCAGCAGGCTCTGTGACTTTGGCCATGTGTGTCTTCAACCTG	1861	2882	GAGCATCGCCACCTGGCATCAGAGCCAGATCCAGCTTCAGCTTCAGCTTCAGCTTCACCA	2941
Db	1803	CCAGGATGGAGCCAGCAGGCTCTGTGACTTTGGCCATGTGTGTCTTCAACCTG	1862	2883	GAGCATCGCCACCTGGCATCAGAGCCAGATCCAGCTTCAGCTTCAGCTTCAGCTTCACCA	2942
QY	1862	ATGGCTTGGAAAGTCTTCTCAAGGGGACTACATCCCTGGCCACAGAGCCACATGG	1921	2942	AAGACGGGACGCTGTTCGCTACGACATGAGAGTGCAGACTCGGGCATTCGACCTGCAGT	3001
Db	1863	ATGGCTTGGAAAGTCTTCTCAAGGGGACTACATCCCTGGCCACAGAGCCACATGG	1922	2943	AAGACGGGACGCTGTTCGCTACGACATGAGAGTGCAGACTCGGGCATTCGACCTGCAGT	3002
QY	1922	CTCGGAGGTGTCTGGGACAGGCTGCGACGCCAAGGTGGATGTCTGGAGCAGTGT	1981	3002	GCACATGGCCCCCTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGCTGG	3061
Db	1923	CTCGGAGGTGTCTGGGACAGGCTGCGACGCCAAGGTGGATGTCTGGAGCAGTGT	1982	3003	GCACATGGCCCCCTGATGGCAGCTTCGCTCGAGCTGGAGGTCAAGCATGGCCAGCTGG	3062
QY	1982	GTATGATGTGCATCATCTCAAGGGTGCACCCCTGGACTCAGTCTTCCGAGGGCCGC	2041	3062	AGAACAGGCCCCCTAACCTTCGCTCCACCGCGGCTCCACATTCGCGG-AAAGCAGCTTC	3121
Db	1983	GTATGATGTGCATCATCTCAAGGGTGCACCCCTGGACTCAGTCTTCCGAGGGCCGC	2042	3063	AGAACAGGCCCCCTAACCTTCGCTCCACCGCGGCTCCACATTCGCGG-AAAGCAGCTTC	3122
QY	2042	TCTGCCTCAAGATTGCCAGAGCCTCGCTGTGAGGAGATCCACCTCTGCGGCC	2101	3122	TGCTCGGTGACAGTGTCTGCTGAAAACACAGGCTCAGCCGTTCCAGGGGATCTGCCAG	3181
Db	2043	TCTGCCTCAAGATTGCCAGAGCCTCGCTGTGAGGAGATCCACCTCTGCGGCC	2102	3122	TGCTCGGTGACAGTGTCTGCTGAAAACACAGGCTCAGCCGTTCCAGGGGATCTGCCAG	3180
QY	2102	CTCTCAGAGCCAGGCTCAAGAGGGCTGAGGAAAGGCCCATCCACCGGGTGTCTG	2161	3182	CCCCCGGCTCAGTGTGGAAACAGGCTTCGACAGCAAGGTGGGGGCAAGAGAAATG	3241
Db	2103	CTCTCAGAGCCAGGCTCAAGAGGGCTGAGGAAAGGCCCATCCACCGGGTGTCTG	2162	3181	-CCCCCGGCTCAGTGTGGAAACAGGCTTCGACAGCAAGGTGGGGGCAAGAGAAATG	3239
QY	2162	CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAGAGCC	2221	3242	CCTCCAGGATTTTACACCTGAGCCCTGCCCACTGCTGAAAAAACH-TCCGCCAGCT	3300
Db	2163	CAGCGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAGAGCC	2222	3240	CCTCCAGGATTTTACACCTGAGCCCTGCCCACTGCTGAAAAAACH-TCCGCCAGCT	3299
QY	2222	CTTGGAGGAGATATTAAGAAACAGACATCCACCGCCAAATCAAGGCCAATTTACCAC	2281	3301	GAAGAGACAGAGAGGATGGCAGGATTTTACCTGGGAAACAAAACAGGATCTTTTCT	3360
Db	2223	CTTGGAGGAGATATTAAGAAACAGACATCCACCGCCAAATCAAGGCCAATTTACCAC	2282	3300	GAAGAGACAG-AGGAGGATGGCAGGATTTTACCTGGGAAACAAAACAGG---ATCTTCTC	3354
QY				3361	GCCTCTGCTCCAGTTCGAGTTGGCTTGCCTTGCCTTGGATCAGTGCATTTGTTGGCAGAC	3420



Db 3355 TGCCCTGCTCCAGTCCAGTTGGCTTGACCCCGTTGGATCAGTGACCAATTTGCTGCGACAC 3414  
Qy 3421 AGGGAGAGCAGCTTTCCAGCCCTGGGTGAGAGGGGTGGCGAGCCCTTCGGGCCCTTCACC 3480  
Db 3415 AGGGAGAGCAGCTTTCCAGCCCTGGGTGAGAGGGGTGGCGAGCCCTTCGGGCCCTTCACC 3474  
Qy 3481 CTCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAACTCAGAGTTTCAGTGCAGAA 3540  
Db 3475 CTCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAACTCAGAGTTTCAGTGCAGAA 3534  
Qy 3541 CCAGGTACAGAGTATGCGCCCGCTGAGTTAAAGGGGCGCTCTAAACCCCTTCGCTGGC 3600  
Db 3535 CCAGGTACAGAGTATGCGCCCGCTGAGTTAAAGGGGCGCTCTAAACCCCTTCGCTGGC 3594  
Qy 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAACTCCCTGACCCCTGGGAAG 3660  
Db 3595 CTCACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAACTCCCTGACCCCTGGGAAG 3654  
Qy 3661 GCTCCTGGTAAATACACACACTTTTCAGAGTTGTTCGAAACACAGGTCCTGAGTTGACC 3720  
Db 3655 GCTCCTGGTAAATACACACACTTTTCAGAGTTGTTCGAAACACAGGTCCTGAGTTGACC 3714  
Qy 3721 TCTGTTCCAGCAAGGACCAAGAGGTGCTAAGTGAAGTGTCTCAGTCCCGACACA 3780  
Db 3715 TCTGTTCCAGCAAGGACCAAGAGGTGCTAAGTGAAGTGTCTCAGTCCCGACACA 3774  
Qy 3781 TGTGCCCCCTTGTGCTGGCTACCACTCTTCCCGACAGCAGCCCGGAGCCCTTCCTCA 3840  
Db 3775 TGTGCCCCCTTGTGCTGGCTACCACTCTTCCCGACAGCAGCCCGGAGCCCTTCCTCA 3834  
Qy 3841 GGCCAGCAGCTGCCCCAGACTCGTGGCACTCAGTTCCTCATCTGTAAGGTTGAAGGTT 3900  
Db 3835 GGCCAGCAGCTGCCCCAGACTCGTGGCACTCAGTTCCTCATCTGTAAGGTTGAAGGTT 3894  
Qy 3901 GATCAGAGATGCTGTACAGACAGTCTGTGATGATGATGATGATGATGATGATGATGATG 3960  
Db 3895 GATCAGAGATGCTGTACAGACAGTCTGTGATGATGATGATGATGATGATGATGATGATG 3954  
Qy 3961 CAGCAGAGAGAGCG-TCCGGCGCCAGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCC 4019  
Db 3955 CAGCAGAGAGAGCGTCCGGCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCC 4014  
Qy 4020 CCAGAGCAGCTCAGCATCACTGACACTCACCTGCGCTGCGCTGCGCCGACAGGGTA 4079  
Db 4015 CCAGAGCAGCTCAGCATCACTGACACTCACCTGCGCTGCGCTGCGCCGACAGGGTA 4074  
Qy 4080 CTGCGGACGGCACTTTGCACTCTGATGACCTCAAGACATTTTCATGGTGCCCTCTGGCA 4139  
Db 4075 CTGCGGACGGCACTTTGCACTCTGATGACCTCAAGACATTTTCATGGTGCCCTCTGGCA 4134  
Qy 4140 GGGCAGGCGAGGCGAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGA 4199  
Db 4135 GGGCAGGCGAGGCGAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGA 4194  
Qy 4200 CACAGTCTTGAGCTGTCCACATGATGATGATCTCTCAAACTCTTCCAGATTTCTTAAG 4259  
Db 4195 CACAGTCTTGAGCTGTCCACATGATGATGATCTCTCAAACTCTTCCAGATTTCTTAAG 4254  
Qy 4260 AATAGCACCCCTTCCCGATGTCCTGAGTCTGCTCTTCCAGGGGAGTACTCAGG 4319  
Db 4255 AATAGCACCCCTTCCCGATGTCCTGAGTCTGCTCTTCCAGGGGAGTACTCAGG 4314  
Qy 4320 ACTACCTAGCATTAATACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 4379  
Db 4315 ACTACCTAGCATTAATACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATG 4374  
Qy 4380 TGGGGCAGGGGACCCCGAGCTCCGCTGGGAGAACTCATTTCCCACTCTTGCCAAAGACAG 4439  
Db 4375 TGGGGCAGGGGACCCCGAGCTCCATGGGAGAACTCATTTCCCACTCTTGCCAAAGACAG 4434  
Qy 4440 CCTTTGTCCAGCTGTCCACATTTGAGTCACTGCTCCCGGGGAGAGCCCGGCCCCCA 4499

Db 4435 CCTTTGTCCAGCTGTCCACATTTGAGTCACTGCTCCCGGGGAGAGAGCCCGGCCCCCA 4494  
Qy 4500 GCACATAAAGAACTGCAGCCCTTGTACTGTCAGAGTCTCGGGTTGTAGAGAACTCTTTGTAA 4559  
Db 4495 GCACATAAAGAACTGCAGCCCTTGTACTGTCAGAGTCTCGGGTTGTAGAGAACTCTTTGTAA 4554  
Qy 4560 GCAATAAAGTTTGGGGTGATGACAAATGTT 4589  
Db 4555 GCAATAAAGTTTGGGGTGATGACAAATGTT 4584

RESULT 6  
BC035576  
LOCUS  
DEFINITION Homo sapiens mitogen-activated protein kinase kinase 14,  
mRNA (cDNA clone MGC:45335 IMAGE:5497185), complete cds.  
ACCESSION BC035576  
VERSION BC035576.1 GI:23272579  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 4452)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.D.,  
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
Kernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Warley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.M., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
TITLE Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
JOURNAL MEDLINE  
PUBMED 2388257  
REFERENCE 2 (bases 1 to 4452)  
Strausberg, R.  
Direct Submission  
Submitted (31-JUL-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NTH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-rc@mail.nih.gov](mailto:cgapbs-rc@mail.nih.gov)  
Tissue Procurement: Lou Staudt  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland:  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nhgri.nih.gov](mailto:nisc_mgc@nhgri.nih.gov)  
Akhtar, N., Ayele, N., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tauregeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 59 Row: f Column: 9  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4505396.

## FEATURES

Location/Qualifiers

1..4452  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:45335 IMAGE:5497185"  
/tissue\_type="Lymph, lymphoma"  
/clone\_lib="NIH MGC\_85"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..4452  
/gene="WAP3K14"  
/note="synonyms: NIK, HSN1K, HS, FTD1B"  
/db\_xref="LocusID:9020"  
/db\_xref="MIM:604655"  
71..2914  
/gene="WAP3K14"  
/codon\_start=1  
/product="WAP3K14 protein"  
/protein\_id="AAH35576.1"  
/db\_xref="GI:23272580"  
/db\_xref="LocusID:9020"  
/db\_xref="MIM:604655"  
/translation="MAVMEACPGAGSVAQQKELPKAKETPLGKKQSVVYLEA  
VKSQYVFGCKWEILNDVITKGTAKGSEBAGPAASIIIAQAECSQEFSPFSEFI  
AGSKYSSQESLDQIPNNVAHATEGMARVWKRRSKARKRRKKSSKSLAHAGVA  
LAKPLPRPEQESCTIPVEDBESPLGAPVVRNTPOFTKPLKPGQLCFKQLGELRL  
PALPSELHLISPLQCLNHWKLHHPDQGGPLPETHPPYSRLPHFPPLQPPWK  
PHESEFGLKACVDSQPLPDPLHSLKACVDSQPLPGLHLEPCLSRGAHEKFSVE  
EYLVAHQSSVSGAHSLSLAKTAAARGSRSEPSKTEDENEGVLITKLPVDSY  
YREYVHAHQRLGRGSGFGEVHRMEDQTPQCAVKVRLEVFRAELIMACAGLTSY  
RIVPLGVAREGPWNITMELEGGSLGQVKEQCLPEDRALYVLGQALEGLVLS  
RRLHGDVADNVLLSSDGHAAALCDFGHVCLQPDGLGKSLITGDIYLPGETHWAP  
VVLGRSDAKVDSSCMHLMLGCHPWTQVFRGPLCLKIASPPVREIIPSCAP  
LTAQAIQGLRKPEIHRVSAELGGKVNRLQQVGLKSPWRGEYKPEPRHPPPNQAN  
HQTLPQRELSPRAPRPARETTGRAPKLQPLPPLPPPPPKSPPLTLSEESGMV  
EPLLSLEPAPRAPSPERKATVPEQLQLELEFLNLSQPSFSLSEBQILSCL  
STDLSLSDSKSPKASOSSRDTLSSSVHSSWSOARSSSNMVLARGPTDTPS  
YNGVKVQIQSLNGELHIREFHRVKVGDIGTGLSSQIPAAAFSLVTKDGPVRVDM  
VPDSGLDQCTLPAGPSFAWSRVRXHGQENRP"  
1286..2011  
/gene="WAP3K14"  
/note="S\_TKc; Region: Serine/Threonine protein kinases,  
catalytic domain"  
/db\_xref="CDD:smart00220"

misc\_feature

## ORIGIN

Query Match 92.5%; Score 4249.4; DB 9; Length 4452;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4354; Conservative 0; Mismatches 41; Indels 10; Gaps 7;  
QY 194 GTGTTTGTTCTCTCTCAGGATGAGCACAAAGCCTGGGAGATGGCAGTGATGGAAATGGCCT 253  
DB 33 GAGATCGGCGCTTGGCGCGATGAGCACAAAGCCTGGGAGATGGCAGTGATGGAAATGGCCT 92  
QY 254 GCCAGGTCCCTGGCTCAGCAGTGGGCGACGAGGAACCTCCCAAGCCAAAGGAGA 313  
DB 93 GCCAGGTCCCTGGCTCAGCAGTGGGCGACGAGGAACCTCCCAAGCCAAAGGAGA 152  
QY 314 AGACGCGCCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGAGGCGGTGGAGAGA 373  
DB 153 AGACGCGCCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGAGGCGGTGGAGAGA 212  
QY 374 GCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATACCAAGGGGCACAGCCA 433  
DB 213 GCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATACCAAGGGGCACAGCCA 272

Db 1353 TCAAAAAGTGGCTGGAAGTATTTTCGGGACAGAGGAGCTGATGGCATGTGCAGATGGA 1412  
Qy 1574 CCTCACCAGAAATGTCTCCCTTTGTATGTAGAGCTGTGAGAGAGGGCTTTGGGTCAACATCT 1633  
Db 1413 CCTCACCAGAAATGTCTCCCTTTGTATGTAGAGCTGTGAGAGAGGGCTTTGGGTCAACATCT 1472  
Qy 1634 TCATGGAGCTCTGTGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGAGAGAGGGCTGTCTCC 1693  
Db 1473 TCATGGAGCTCTGTGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGAGAGAGGGCTGTCTCC 1532  
Qy 1694 CAGAGGACCGGGCCCTGTACTACCTGGCCAGGSCCTGGAGGGTCTGGAATACCTCCACT 1753  
Db 1533 CAGAGGACCGGGCCCTGTACTACCTGGCCAGGSCCTGGAGGGTCTGGAATACCTCCACT 1592  
Qy 1754 CACGAAGGATTTCTGATGGGACGCTCAAGCTGACCAACGCTCTCTGTTCAGAGCATGGGA 1813  
Db 1593 CACGAAGGATTTCTGATGGGACGCTCAAGCTGACCAACGCTCTCTGTTCAGAGCATGGGA 1652  
Qy 1814 GCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGATGGCTGGGAA 1873  
Db 1653 GCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTTCAACCTGATGGCTGGGAA 1712  
Qy 1874 AGTCTTGTCTACAGGGGACTACATCCCTTGGCCACAGAGACCCACATGGCTCCGAGGTGG 1933  
Db 1713 AGTCTTGTCTACAGGGGACTACATCCCTTGGCCACAGAGACCCACATGGCTCCGAGGTGG 1772  
Qy 1934 TGCTGGGACGAGCTGCGACGCCCAAGGTGGATGTCTGGAGCAGCTGTGTATGATGTGC 1993  
Db 1773 TGCTGGGACGAGCTGCGACGCCCAAGGTGGATGTCTGGAGCAGCTGTGTATGATGTGC 1832  
Qy 1994 ACATGCTCAACGGGTGCCACCCCTGGACTCAAGTTCTTCCGAGGGCCGCTCTGCCCTCAAGA 2053  
Db 1833 ACATGCTCAACGGGTGCCACCCCTGGACTCAAGTTCTTCCGAGGGCCGCTCTGCCCTCAAGA 1892  
Qy 2054 TTGCCAGGGACCTTCGGCTGTGAGGGAGATCCACCCCTCTGCGCCCTCTCAAGCCC 2113  
Db 1893 TTGCCAGGGACCTTCGGCTGTGAGGGAGATCCACCCCTCTGCGCCCTCTCAAGCCC 1952  
Qy 2114 AGGCCATCAAGAGGGGTGAGGAAGAGGCCATCCACCCGCTGTCTGCGAGCGAGCTGG 2173  
Db 1953 AGGCCATCAAGAGGGGTGAGGAAGAGGCCATCCACCCGCTGTCTGCGAGCGAGCTGG 2012  
Qy 2174 GAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTTGGAGGGAG 2233  
Db 2013 GAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTTGGAGGGAG 2072  
Qy 2234 AATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACACAGACCCTCCATG 2293  
Db 2073 AATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACACAGACCCTCCATG 2132  
Qy 2294 CCCAGCCGAGAGCTTTCCGACAGGGCCCGAGGGCCCGCCAGCTGAGGAGACACAG 2353  
Db 2133 CCCAGCCGAGAGCTTTCCGACAGGGCCCGAGGGCCCGCCAGCTGAGGAGACACAG 2192  
Qy 2354 GCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCGCCAGAGCAAAAGTCTC 2413  
Db 2193 GCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACAGAGCCCGCCAGAGCAAAAGTCTC 2252  
Qy 2414 CTCCCTTGAATTTGAGCAAGGAGGTCTGGGATGTGGGAACCTTACCTCTGTCCTCCC 2473  
Db 2253 CTCCCTTGAATTTGAGCAAGGAGGTCTGGGATGTGGGAACCTTACCTCTGTCCTCCC 2312  
Qy 2474 TGGAGCCAGCCCTTCCGACAGAACCCAGCTCACAGAGCGGAAGACCGTCCCGAGC 2533  
Db 2313 TGGAGCCAGCCCTTCCGACAGAACCCAGCTCACAGAGCGGAAGACCGTCCCGAGC 2372  
Qy 2534 AGGAACGTGACAGCTGGAATAGAAATTAATTCCTCAACAGGCTGTCTCCAGGCAATTTTCTC 2593  
Db 2373 AGGAACGTGACAGCTGGAATAGAAATTAATTCCTCAACAGGCTGTCTCCAGGCAATTTTCTC 2432  
Qy 2594 TGGAGGACGAGGACAAATTTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTGCGATG 2653

Db 2433 TGGAGGACGAGGACAAATTTCTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTGCGATG 2492  
Qy 2654 ACAGTGAAGAAAGACCCATCAAAAGCCCTCTCAAAAGCTCGCGGACACCCCTGAGCTCAGCG 2713  
Db 2493 ACAGTGAAGAAAGACCCATCAAAAGCCCTCTCAAAAGCTCGCGGACACCCCTGAGCTCAGCG 2552  
Qy 2714 TACACTCTGAGAGCAGCCAGGCTCGAAGCTCGAAGCTCAGCTGGAACATGCTGTGTCGCC 2773  
Db 2553 TACACTCTGAGAGCAGCCAGGCTCGAAGCTCGAAGCTCAGCTGGAACATGCTGTGTCGCC 2612  
Qy 2774 GGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCAAAATACAGTCTC 2833  
Db 2613 GGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCAAAATACAGTCTC 2672  
Qy 2834 TTAATGTGAACACCTTCACATCCGGAGTTTCCACCGGGTCAAAAGTGGGAGACATCGCCA 2893  
Db 2673 TTAATGTGAACACCTTCACATCCGGAGTTTCCACCGGGTCAAAAGTGGGAGACATCGCCA 2732  
Qy 2894 CTGSCATCAGAGCAGATCCCAAGCTCAGCTTTCAGCTTGGTTCACCAAGAGCGGGCAGC 2953  
Db 2733 CTGSCATCAGAGCAGATCCCAAGCTCAGCTTTCAGCTTGGTTCACCAAGAGCGGGCAGC 2792  
Qy 2954 CTGTTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGACATGACATCTGGCCC 3013  
Db 2793 CTGTTGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGACATGACATCTGGCCC 2852  
Qy 3014 CTGATGSCAGCTTCGCGCTGGAGGTGGAAGGTCAAGCATGCGCAGCTGGAGAACAGGCCCT 3073  
Db 2853 CTGATGSCAGCTTCGCGCTGGAGGTGGAAGGTCAAGCATGCGCAGCTGGAGAACAGGCCCT 2912  
Qy 3074 AACCTGCTCCACCGCCGGCTCCACACTGCGCGGAAAGCAGCTTCTGCTCGGTGAC 3133  
Db 2913 AACCTGCTCCACCGCCGGCTCCACACTGCGCGG - AAGCAGCCCTCTCTGCTCGGTGAC 2971  
Qy 3134 GATGCTGCTGAAACACAGGCTCAGCCCTTCCAGGGATTCGCGAGCCCGCCGCTCA 3193  
Db 2972 GATGCTGCTG - AAACACAGGCTCAGCCGTTCCAGGGATCTGCCAG - CCCCCGGCTCA 3029  
Qy 3194 CAGTGGGAACAGGGCCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCCTCCAGGAT 3253  
Db 3030 GCAGTGGGACAGGGCCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCCTCCAGGAT 3089  
Qy 3254 TCACACTGAGCCCTGCCCCACCTGTGTGAAAAACA - TCCGCCACTGGAAGACAGAA 3312  
Db 3090 TCACACTGAGCCCTGCCCCACCTGTGTGAAAAACA - TCCGCCACTGGAAGACAG - A 3148  
Qy 3313 GAGGATGCGAGAGTTTACCTGGGAAACAAACAGGGATCTTTTCTGCCCCCTGCTCCA 3372  
Db 3149 GAGGATGCGAGGAGTTTACCTCGGAAACAAACAGG - - - - - ATCTTCTCTGCCCCCTGCTCCA 3204  
Qy 3373 GTGAGTTGGCCCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGACAGGGAGAGCAG 3432  
Db 3205 GTGAGTTGGCCCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGACAGGGAGAGCAG 3264  
Qy 3433 CTTCCAGCCCTGGGTTCAGAAAGGGGTGGCGAGCCCTTCGGGCCCTCACCCCTCAGGCTGCT 3492  
Db 3265 CTTCCAGCCCTGGGTTCAGAAAGGGGTGGCGAGCCCTTCGGGCCCTCACCCCTCAGGCTGCT 3324  
Qy 3493 GTGAGAGTGCAGAGTGTAAAGGGCCCAAACTCAGGTTTCAGTGAGAGAACAGGTCAGGAG 3552  
Db 3325 GTGAGAGTGCAGAGTGTAAAGGGCCCAAACTCAGGTTTCAGTGAGAGAACAGGTCAGGAG 3384  
Qy 3553 GTATGCCCGCCCTGAGGTTAAAGGGGCCCTTAAACCCCTTGCCTGCGCTCACCTGSCCA 3612  
Db 3385 GTATGCCCGCCCTGAGGTTAAAGGGGCCCTTAAACCCCTTGCCTGCGCTCACCTGSCCA 3444  
Qy 3613 GCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCTTGGGAAGGCTCCCTGCTAG 3672  
Db 3445 GCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCTTGGGAAGGCTCCCTGCTAG 3504  
Qy 3673 AATAACACACATTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGTTGTCAGCC 3732  
Db 3505 AATAACACACATTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGTTGTCAGCC 3564

```
QY 3733 AAGGACCAAGAGGTGTGAAGTGGTTCTCAGTCCCGACACATGTGCCCTTTG 3792
Db 3565 AAGGACCAAGAGGTGTGAAGTGGTTCTCAGTCCCGACACATGTGCCCTTTG 3624
QY 3793 CTGTGGCTTACCACCTTTCCCGACAGCAGCGCCCGAGGCCCTTCCAGGCCCGACGACTG 3852
Db 3625 CTGTGGCTTACCACCTTTCCCGACAGCAGCGCCCGAGGCCCTTCCAGGCCCGACGACTG 3684
QY 3853 CCCGAGACTCGTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGTCAGGATAT 3912
Db 3685 CCCGAGACTCGTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGTCAGGATAT 3744
QY 3913 GCCTCAGAGGAACAGTCTGTGATGAGACATGATCAGTCTAAGGAAACAGCAGAGAGAG 3972
Db 3745 GCCTCAGAGGAACAGTCTGTGATGAGACATGATCAGTCTAAGGAAACAGCAGAGAGAG 3804
QY 3973 AGC-TCCGGCGCCCGACGCCCACTATCAGTGTCCAGCGTGTGGTTCCCGAGACACAGC 4031
Db 3805 AGCTCCGGCGCCCGACGCCCACTATCAGTGTCCAGCGTGTGGTTCCCGAGACACAGC 3864
QY 4032 TCAGCATCACATGACATCACCTGCCCTGCCCTGGCCAGAGGGTACTGCCGACGCA 4091
Db 3865 TCAGCATCACATGACATCACCTGCCCTGCCCTGGCCAGAGGGTACTGCCGACGCA 3924
QY 4092 TTTTGCACCTCTGATGACCTCAAAGCACATTTTCATGTGCTGCCCTCTGGCAGGGCAGGCAGG 4151
Db 3925 CTTTGCACCTCTGATGACCTCAAAGCACATTTTCATGTGCTGCCCTCTGGCAGGGCAGGGTAGG 3984
QY 4152 GCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGAACACAGTCTTGAG 4211
Db 3985 GCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGAACACAGTCTTGAG 4044
QY 4212 CTGTCCACATGATGTGACTCTCAAACCTCTTCCAGATTTCTTAAGATAGCACCCCC 4271
Db 4045 CTGTCCACATGATGTGACTCTCAAACCTCTTCCAGATTTCTTAAGATAGCACCCCC 4104
QY 4272 TTCCCCATTGCCCGAGCTTAGCCTCTTCCAGAGGAGCTACTCAGGACTCACGTAGCA 4331
Db 4105 TTCCCCATTGCCCGAGCTTAGCCTCTTCCAGAGGAGCTACTCAGGACTCACGTAGCA 4164
QY 4332 TTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCTCGGGGCGAGGGA 4391
Db 4165 TTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCTCGGGGCGAGGGA 4224
QY 4392 CGCCGAGACTCGTGGGAGAAGCTCATTCACATCTTGCAGAGAGCGCTTTGTCCAGC 4451
Db 4225 CGCCGAGACTCGTGGGAGAAGCTCATTCACATCTTGCAGAGAGCGCTTTGTCCAGC 4284
QY 4452 TGTCCACATTGAGTCAGACTGCTCCCGGGGAGAGAGCCCGCCCGCCAGCACATAAGAA 4511
Db 4285 TGTCCACATTGAGTCAGACTGCTCCCGGGGAGAGAGCCCGCCCGCCAGCACATAAGAA 4344
QY 4512 CTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTTTTGTAAAGCAATAAGTTT 4571
Db 4345 CTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTTTTGTAAAGCAATAAGTTT 4404
QY 4572 GGGGTGATCACAAATGTTAAAAAAA 4596
Db 4405 GGGGTGATCACAAATGTTAAAAAAA 4429
```

```
RESULT 7
AR062290 LOCUS AR062290 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5843721.
ACCESSION AR062290
VERSION AR062290.1 GI:5989981
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3156)
```

AUTHORS Rothe,M. and Wu,L.  
TITLE Nucleic acids encoding human NIK protein  
JOURNAL Patent: US 5843721-A 1 01-DEC-1998;  
FEATURES Location/Qualifiers  
source 1..3156  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

```
Query Match 68.5%; Score 3149.6; DB 6; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 232 ATGCGAGTGTATGGAATGGCTGTGCCAGGTGCCCTCGCTCAGCAGTGGGSCAGCAGAAG 291  
Db 1 ATGCGAGTGTATGGAATGGCTGTGCCAGGTGCCCTCGCTCAGCAGTGGGSCAGCAGAAG 60  
QY 292 GAACTCCCAAGCCAAAGGAGAGACGCCGCCACTGGGGAAGAAACAGAGCTCGTCTAC 351  
Db 61 GAACTCCCAAGCCAAAGGAGAGACGCCGCCACTGGGGAAGAAACAGAGCTCGTCTAC 120  
QY 352 AAGCTTGGGCGCGTGGAGAGAGAGCCCTGTCTCGCGAAAGTGGGAGATCCTCAATGAC 411  
Db 121 AAGCTTGGGCGCGTGGAGAGAGAGCCCTGTCTCGCGAAAGTGGGAGATCCTCAATGAC 180  
QY 412 GTGATTACCAAGGCGCAGCCCAAGGAAGGCTCCGAGGCAGGCGCAGCTGCCATCTCTATC 471  
Db 181 GTGATTACCAAGGCGCAGCCCAAGGAAGGCTCCGAGGCAGGCGCAGCTGCCATCTCTATC 240  
QY 472 ATCGCCGAGGCTGAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAGCGCAT 531  
Db 241 ATCGCCGAGGCTGAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAGCGCAT 300  
QY 532 TTCAATCGTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 591  
Db 301 TTCAATCGTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 360  
QY 592 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGGAAAGGGAAGACGTGCGACG 651  
Db 361 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGGAAAGGGAAGACGTGCGACG 420  
QY 652 AAAGCCCGGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGCGAGGATGGCC 711  
Db 421 AAAGCCCGGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGCGAGGATGGCC 480  
QY 712 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 771  
Db 481 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGAGAGCTGCACCATCCAGTGCAGGAG 540  
QY 772 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCCACCAAGCCTCTG 831  
Db 541 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCCACCAAGCCTCTG 600  
QY 832 AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCT 891  
Db 601 AAGGAACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCT 660  
QY 892 CTGCTCGATCAGAACTCCACAACTGATCAGCCCTTGCATATGTCGAACACACGTGTGG 951  
Db 661 CTGCTCGATCAGAACTCCACAACTGATCAGCCCTTGCATATGTCGAACACACGTGTGG 720  
QY 952 AAATGCAACACCCCGAGGAGGCGCCCTTGCCCTCGCCACGACCCCTTCCCTTAT 1011  
Db 721 AAATGCAACACCCCGAGGAGGCGCCCTTGCCCTCGCCACGACCCCTTCCCTTAT 780  
QY 1012 AGCAGACTGCTCATCTCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCACCTCTG 1071  
Db 781 AGCAGACTGCTCATCTCCCTTCCATTCACCTCTCCAGCCCTGGAACCTCACCTCTG 840  
QY 1072 GAGTCTTCTGGGCAAACTGCGCTGTGTAGACAGCCAGAAACCTTGGCTCAGCCACAC 1131  
Db 841 GAGTCTTCTGGGCAAACTGCGCTGTGTAGACAGCCAGAAACCTTGGCTCAGCCACAC 900
```

QY 1132 CTGACCAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGTGAGGCC 1191  
DB 901 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGTGAGGCC 960  
QY 1192 AGCTGCCCTGTCTCGTGGTGGCCCATGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1251  
DB 961 AGCTGCCCTGTCTCGTGGTGGCCCATGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1020  
QY 1252 CTGCAAGGAGCGGTGAGTCTCAAGCCAGCCCAACAGCCTGACAGCCTGGCCCAAGACCTGG 1311  
DB 1021 CTGCAAGGAGCGGTGAGTCTCAAGCCAGCCCAACAGCCTGACAGCCTGGCCCAAGACCTGG 1080  
QY 1312 GCAGACCGGGCTCCAGATCCCGGAGCCGAGCCCAAACTGAGGACAAAGAGGGTGTCTC 1371  
DB 1081 GCAGACCGGGCTCCAGATCCCGGAGCCGAGCCCAAACTGAGGACAAAGAGGGTGTCTC 1140  
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATAGTACCGAGAAAGTCCACTGGGCC 1431  
DB 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTATAGTACCGAGAAAGTCCACTGGGCC 1200  
QY 1432 ACGCACCAAGCTCCGCTCGGAGAGGCTCTTTCGGAGAGGTGCAAGGATGGAGGACAAG 1491  
DB 1201 ACGCACCAAGCTCCGCTCGGAGAGGCTCTTTCGGAGAGGTGCAAGGATGGAGGACAAG 1260  
QY 1492 CAGACTGGCTTCCAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTCCGGGAGAGGAG 1551  
DB 1261 CAGACTGGCTTCCAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTCCGGGAGAGGAG 1320  
QY 1552 CTGATGGCATGTGACAGATTGACCTCACCAAGATTGTCCCTTTGTATGGAGCTGTGAGA 1611  
DB 1321 CTGATGGCATGTGACAGATTGACCTCACCAAGATTGTCCCTTTGTATGGAGCTGTGAGA 1380  
QY 1612 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671  
DB 1381 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440  
QY 1672 GTCAAGGAGCAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1731  
DB 1441 GTCAAGGAGCAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1500  
QY 1732 GAGGCTCTGGAATACCTCCACTCAGAAAGATTCTGCATGGGGAGCTCAAGCTGACAAAC 1791  
DB 1501 GAGGCTCTGGAATACCTCCACTCAGAAAGATTCTGCATGGGGAGCTCAAGCTGACAAAC 1560  
QY 1792 GTGCTCTGTCTCAGCGATGGAGGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851  
DB 1561 GTGCTCTGTCTCAGCGATGGAGGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620  
QY 1852 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAG 1911  
DB 1621 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAG 1680  
QY 1912 ACCCAGATGGCTCCGAGGTGGTGTGGGAGGAGCTGCAGACGCAAGGTGGATGTCTGG 1971  
DB 1681 ACCCAGATGGCTCCGAGGTGGTGTGGGAGGAGCTGCAGACGCAAGGTGGATGTCTGG 1740  
QY 1972 AGCAGCTGCTGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTCTC 2031  
DB 1741 AGCAGCTGCTGTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTCTC 1800  
QY 2032 CGAGGGCGCTCTGCCTCAAGATTGCCAGAGCCTCCGCTGTGAGGGAGATCCCAACC 2091  
DB 1801 CGAGGGCGCTCTGCCTCAAGATTGCCAGAGCCTCCGCTGTGAGGGAGATCCCAACC 1860  
QY 2092 TCCTGGCCCTCTTCAGGCCAGGCCATCAAGAGGGGCTGAGAAAGAGCCCATCCAC 2151  
DB 1861 TCCTGGCCCTCTTCAGGCCAGGCCATCAAGAGGGGCTGAGAAAGAGCCCATCCAC 1920  
QY 2152 CGCGTGTCTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGT 2211  
DB 1921 CGCGTGTCTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGT 1980  
QY 2212 CTGAAGAGCCCTTGGAGGGGAGATATATAAGAACCAAGACATCCACCCGCAATCAAGCC 2271

DB 1981 CTGAAGAGCCCTTGGAGGGGAGATATATAAGAACCAAGACATCCACCCCAATCAAGCC 2040  
QY 2272 AATTACCAACAGACCCCTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2331  
DB 2041 AATTACCAACAGACCCCTCATGCCAGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2100  
QY 2332 CGGCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCCAACAGAG 2391  
DB 2101 CGGCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCCAACAGAG 2160  
QY 2392 CCCCCAGAGCCAAACAGTCTCTCCCTTGAATTTGAGCAAGAGGAGTCTGGGATGTGG 2451  
DB 2161 CCCCCAGAGCCAAACAGTCTCTCCCTTGAATTTGAGCAAGAGGAGTCTGGGATGTGG 2220  
QY 2452 GAAACCTTACCTCTGTCTCTCGAGCCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2511  
DB 2221 GAAACCTTACCTCTGTCTCTCGAGCCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2280  
QY 2512 CGGAAAGCAACCCGTCGAGAGCACTGTCAGAGCTGGAAATAGAAATTTCTCTCAAC 2571  
DB 2281 CGGAAAGCAACCCGTCGAGAGCACTGTCAGAGCTGGAAATAGAAATTTCTCTCAAC 2340  
QY 2572 AGCCTGTCCCAGCCATTTTCTCTGAGAGCAGAGCAAAATCTCTCGTGTCTCAGATC 2631  
DB 2341 AGCCTGTCCCAGCCATTTTCTCTGAGAGCAGAGCAAAATCTCTCGTGTCTCAGATC 2400  
QY 2632 GACAGCTCTCTCTCGGATGACAGTGAGAAAGAACCCATCAAGGGCTCTCAAGAGCTCG 2691  
DB 2401 GACAGCTCTCTCTCGGATGACAGTGAGAAAGAACCCATCAAGGGCTCTCAAGAGCTCG 2460  
QY 2692 CGGAGACCCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGGCTCGAAGCTCC 2751  
DB 2461 CGGAGACCCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCCAGGGCTCGAAGCTCC 2520  
QY 2752 AGCTGGAAATAGTGTCTGGCCCGGGGGGGCCCAACGACACCCCAAGCTATTTCAATGGT 2811  
DB 2521 AGCTGGAAATAGTGTCTGGCCCGGGGGGGCCCAACGACACCCCAAGCTATTTCAATGGT 2580  
QY 2812 GTGAAGTCCAAATACAGTCTCTTAATGTGTGAACACCTGCACATCCGGGAGTTCACCCGG 2871  
DB 2581 GTGAAGTCCAAATACAGTCTCTTAATGTGTGAACACCTGCACATCCGGGAGTTCACCCGG 2640  
QY 2872 GTCAAAAGTGGGAGACATCGCCACTGGCATCAGACGACAGATCCAGCTGCAAGCTTCAGC 2931  
DB 2641 GTCAAAAGTGGGAGACATCGCCACTGGCATCAGACGACAGATCCAGCTGCAAGCTTCAGC 2700  
QY 2932 TTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGACATGGAGGTGCCAGACTCGGGCATC 2991  
DB 2701 TTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGACATGGAGGTGCCAGACTCGGGCATC 2760  
QY 2992 GACCTGAGTGCAACATCGGCCCTTATGGAGCTTCGCTGGAGCTGGAGGCTCAAGCAT 3051  
DB 2761 GACCTGAGTGCAACATCGGCCCTTATGGAGCTTCGCTGGAGCTGGAGGCTCAAGCAT 2820  
QY 3052 GGCAGCTGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCGGAAA 3111  
DB 2821 GGCAGCTGGAGAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCGGAAA 2880  
QY 3112 GCAGCTTCTGTCTGTGTGACGATGTGCTGTGCTGAAACACAGGCTCAGCGTTCACAGG 3171  
DB 2881 GCAGCTTCTGTCTGTGTGACGATGTGCTGTGCTGAAACACAGGCTCAGCGTTCACAGG 2940  
QY 3172 GGATTCAGCCCGCCCGGCTCAGTGGGAAACAGGGCTTCGAGAGCAGAGGTGGGGGC 3231  
DB 2941 GGATTCAGCCCGCCCGGCTCAGTGGGAAACAGGGCTTCGAGAGCAGAGGTGGGGGC 3000  
QY 3232 AAGCAGAAATGCTCCCAAGGATTTTCAACCTGAGCCCTGCCACCTGCTGAAAAAATCAT 3291  
DB 3001 AAGCAGAAATGCTCCCAAGGATTTTCAACCTGAGCCCTGCCACCTGCTGAAAAAATCAT 3060  
QY 3292 CCGCCACGTGAAGAGACAGAGGAGGTGACAGAGTTTACTTGGGGAAACAAAAACAGGGA 3351

Db	3061	CGCCACGCTGAAGACAGAGAGGAGATGGCAGAGTTACTCTGGGAAACAAAACAGGGA	3120	
Qy	3352	TCTTTTCTGCCCCCTGCTCCAGTCGAGTTGGCCTGA	3387	
Db	3121	TCTTTTCTGCCCCCTGCTCCAGTCGAGTTGGCCTGA	3156	
RESULT 8				
AR063065				
LOCUS	AR063065	3156 bp	DNA	linear
DEFINITION	Sequence 1 from patent US 5844073.			
ACCESSION	AR063065			
VERSION	AR063065.1	GI:5990756		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 3156)			
AUTHORS	Rothe,M. and Wu,L.			
TITLE	Human NIK proteins			
JOURNAL	Patent: US 5844073-A 1 01-DEC-1998;			
FEATURES	Location/Qualifiers			
source	1..3156			
	/organism="unknown"			
	/mol_type="unassigned DNA"			
ORIGIN				
	Query Match	68.5%;	Score 3149.6;	DB 6; Length 3156;
	Best Local Similarity	99.9%;	Pred. No. 0;	
	Matches 3152;	Conservative	0; Mismatches	4; Indels 0; Gaps 0;
Qy	232	ATGGCAGTATGGAATGGCCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAG	291	
Db	1	ATGGCAGTATGGAATGGCCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAG	60	
Qy	292	GAACTCCCAAGCCAAAGGAGAAGACGCCCACTGGGGAAGAAAACAGAGTCCGTCTAC	351	
Db	61	GAACTCCCAAGCCAAAGGAGAAGACGCCCACTGGGGAAGAAAACAGAGTCCGTCTAC	120	
Qy	352	AAGCTTGAGCCGTGGAGAGCCCTGTCTCGGAAAGTGGGAGATCCTGAATGAC	411	
Db	121	AAGCTTGAGCCGTGGAGAGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGAC	180	
Qy	412	GTGATTACCAAGGGCACAGCCAAAGGAGGCTCCGAGCAGGGCCAGCTCTCTATC	471	
Db	181	GTGATTACCAAGGGCACAGCCAAAGGAGGCTCCGAGCAGGGCCAGCTCTCTATC	240	
Qy	472	ATCGCCAGGCTGAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAAACGCATT	531	
Db	241	ATCGCCAGGCTGAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAAACGCATT	300	
Qy	532	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAT	591	
Db	301	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAT	360	
Qy	592	GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTCCGACG	651	
Db	361	GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTCCGACG	420	
Qy	652	AAAGCCCGGAAGAAAACGGAAGAGAGAGTCAAAAGTCCCTGGCTCATGAGAGTGGCC	711	
Db	421	AAAGCCCGGAAGAAAACGGAAGAGAGAGTCAAAAGTCCCTGGCTCATGAGAGTGGCC	480	
Qy	712	TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGTGCACCAATCCCAAGTGCAGGAG	771	
Db	481	TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGTGCACCAATCCCAAGTGCAGGAG	540	
Qy	772	GATGAGTCTCCACTCGGCCGCCCATATGTTAGAAACACCCCGAGTTACCAAGCCTCTG	831	
Db	541	GATGAGTCTCCACTCGGCCGCCCATATGTTAGAAACACCCCGAGTTACCAAGCCTCTG	600	
Qy	832	AAGGAACACAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGCCCGGCT	891	

Db	601	AAGGAAACCGAGCCCTTGGGCAAACTCTGTGTTTAAACGAGCTTGGCGAGGCGCTACGCGCGCT	660	
Qy	892	CTGCCTCGATCAGAACTCCACAAACTGATCAGGCCCCCTTGCAATGTCTGAAACCAACGTGTGG	951	
Db	661	CTGCCTCGATCAGAACTCCACAAACTGATCAGGCCCCCTTGCAATGTCTGAAACCAACGTGTGG	720	
Qy	952	AAACTGCACACCCCCCAGGACGGAGGCCCTCGCCCTGCCCCAGCAGCAACCCCTTCCCTAT	1011	
Db	721	AAACTGCACACCCCCCAGGACGGAGGCCCTCGCCCTGCCCCAGCAGCAACCCCTTCCCTAT	780	
Qy	1012	AGCAGACTGCTCATCCCTTCCCATTCACCTCTCCAGCCCTTGGAAACCTCAACCTCTG	1071	
Db	781	AGCAGACTGCTCATCCCTTCCCATTCACCTCTCCAGCCCTTGGAAACCTCAACCTCTG	840	
Qy	1072	GAGTCTCTTCTGGGCAAACTGGCCCTGTGTAGACAGCCAGCAAAACCTTGGCTGACCCACAC	1131	
Db	841	GAGTCTCTTCTGGGCAAACTGGCCCTGTGTAGACAGCCAGCAAAACCTTGGCTGACCCACAC	900	
Qy	1132	CTGAGCAAACTGGCCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGGAGCCC	1191	
Db	901	CTGAGCAAACTGGCCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGGAGCCC	960	
Qy	1192	AGCTGCTGTCTGCTGCTGCCATGAGAGTTTCTGTGAGGGAATACCTAGTGCATGCT	1251	
Db	961	AGCTGCTGTCTGCTGCTGCCATGAGAGTTTCTGTGAGGGAATACCTAGTGCATGCT	1020	
Qy	1252	CTGCAAGCGAGCTGAGCTCAAGCCAGGCCACACAGCTGACAGCTGGCCCAAGACCTGG	1311	
Db	1021	CTGCAAGCGAGCTGAGCTCAAGCCAGGCCACACAGCTGACAGCTGGCCCAAGACCTGG	1080	
Qy	1312	GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCCCAAACTGAGGCAACAGAGGGTGT	1371	
Db	1081	GCAGCAGGGGCTCCAGATCCCGGAGCCAGCCCCCAAACTGAGGCAACAGAGGGTGT	1140	
Qy	1372	CTGCTCTCTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAGAGTCCACTGGGCC	1431	
Db	1141	CTGCTCTCTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAGAGTCCACTGGGCC	1200	
Qy	1432	ACGCACACAGCTCCGCTGGGAGAGGCTCTTCGAGAGGTGCAAGATGGAGACAG	1491	
Db	1201	ACGCACACAGCTCCGCTGGGAGAGGCTCTTCGAGAGGTGCAAGATGGAGACAG	1260	
Qy	1492	CAGACTGGCTTCAGTGGCTGTCAAAGGCTGGCCCTGGAAGTATTTTCGGGCGAGAGAG	1551	
Db	1261	CAGACTGGCTTCAGTGGCTGTCAAAGGCTGGCCCTGGAAGTATTTTCGGGCGAGAGAG	1320	
Qy	1552	CTGATGGCATGTGAGGATTCACCTCAACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGA	1611	
Db	1321	CTGATGGCATGTGAGGATTCACCTCAACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGA	1380	
Qy	1612	GAGGGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG	1671	
Db	1381	GAGGGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG	1440	
Qy	1672	GTCAGGACGAGGCTGTCTCCAGAGACCGGGCCCTGTACTACCTGGGCGAGGCCCTG	1731	
Db	1441	GTCAGGACGAGGCTGTCTCCAGAGACCGGGCCCTGTACTACCTGGGCGAGGCCCTG	1500	
Qy	1732	GAGGGTCTGGAATACCTCACTCAAGAGGATTTCTGCATGGGAGCGTCAAAAGCTCAAC	1791	
Db	1501	GAGGGTCTGGAATACCTCACTCAAGAGGATTTCTGCATGGGAGCGTCAAAAGCTCAAC	1560	
Qy	1792	GTGCTCTCTGTCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGT	1851	
Db	1561	GTGCTCTCTGTCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGT	1620	
Qy	1852	CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGAGGAGTACATCCCTGGCAGAG	1911	
Db	1621	CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGAGGAGTACATCCCTGGCAGAG	1680	
Qy	1912	ACCCACATGGCTCCGAGAGTGTGTGGCAGGAGCTGCGACGCCAAGGTGATGCTGTG	1971	
Db	1681	ACCCACATGGCTCCGAGAGTGTGTGGCAGGAGCTGCGACGCCAAGGTGATGCTGTG	1740	



```
QY 1972 AGCAGCTGTGTATGATGTGACATGCTCAACGGCTGCCACCCCTGGAACCTGACATGAGTTTTC 2031
Db 1741 AGCAGCTGTGTATGATGTGACATGCTCAACGGCTGCCACCCCTGGAACCTGACATGAGTTTTC 1800
QY 2032 CGAGGGCGCTCTGCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCACCC 2091
Db 1801 CGAGGGCGCTCTGCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCACCC 1860
QY 2092 TCCTGCGCCCTCTCACAGCCAGCCCATCAAGAGGGGCTGAGAAAGAGCCCATCCAC 2151
Db 1861 TCCTGCGCCCTCTCACAGCCAGCCCATCAAGAGGGGCTGAGAAAGAGCCCATCCAC 1920
QY 2152 CGCTGTCTGACGCGGAGCTGGAGGAGAGGTGAACCGGGCACTACAGCAAGTGGAGGT 2211
Db 1921 CGCTGTCTGACGCGGAGCTGGAGGAGAGGTGAACCGGGCACTACAGCAAGTGGAGGT 1980
QY 2212 CTGAAGAGCCCTTGAGGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCC 2271
Db 1981 CTGAAGAGCCCTTGAGGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCC 2040
QY 2272 AATTACCAACAGACCTCCATGCCAGCCGAGAGAGCTTTGCCAAAGGGCCCCAGGGGCC 2331
Db 2041 AATTACCAACAGACCTCCATGCCAGCCGAGAGAGCTTTGCCAAAGGGCCCCAGGGGCC 2100
QY 2332 CGGCCAGCTGAGGAGAACACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCCAACAGAG 2391
Db 2101 CGGCCAGCTGAGGAGAACACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCCAACAGAG 2160
QY 2392 CCCCAGAGCAACCAAGTCTCTCCCTTGACTTTGACCAAGGAGGAGTCTGGGATGTGG 2451
Db 2161 CCCCAGAGCAACCAAGTCTCTCCCTTGACTTTGACCAAGGAGGAGTCTGGGATGTGG 2220
QY 2452 GAACCTTACTCTGTCTCTCTGAGCGAGCCCTGCGAGAACCCCGAGCTCAACAGAG 2511
Db 2221 GAACCTTACTCTGTCTCTCTGAGCGAGCCCTGCGAGAACCCCGAGCTCAACAGAG 2280
QY 2512 CGGAAGCAACCGTCCGAGCAGGAACTGAGCAGCTGGAATATGATTTCTCTCAAC 2571
Db 2281 CGGAAGCAACCGTCCGAGCAGGAACTGAGCAGCTGGAATATGATTTCTCTCAAC 2340
QY 2572 AGCTGTCTCCAGCCATTTCTCTGAGAGCAGAGGAGCAATTTCTCTGTGCTCAGCATC 2631
Db 2341 AGCTGTCTCCAGCCATTTCTCTGAGAGCAGAGGAGCAATTTCTCTGTGCTCAGCATC 2400
QY 2632 GACAGCTCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAGGGCTCTCAAGCTCG 2691
Db 2401 GACAGCTCTCCCTGTGCGATGACAGTGAGAGAACCCATCAAGGGCTCTCAAGCTCG 2460
QY 2692 CGGGACACCTTGAGCTCAGGGGTACACTCTCTGAGCAGCAGCGCGAGGCTCGAAGCTCC 2751
Db 2461 CGGGACACCTTGAGCTCAGGGGTACACTCTCTGAGCAGCAGCGCGAGGCTCGAAGCTCC 2520
QY 2752 AGCTGGAAATGTTGCTGGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2811
Db 2521 AGCTGGAAATGTTGCTGGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580
QY 2812 GTGAAGTCCAAATACAGTCTTTAATGGTGAACACCTGCAATCCGGAGTTCCACCGG 2871
Db 2581 GTGAAGTCCAAATACAGTCTTTAATGGTGAACACCTGCAATCCGGAGTTCCACCGG 2640
QY 2872 GTCAAGTGGGAGACATCCCACTGGCATCAGCAGCCAGATCCCAAGCTGCAAGCTTCAAGC 2931
Db 2641 GTCAAGTGGGAGACATCCCACTGGCATCAGCAGCCAGATCCCAAGCTGCAAGCTTCAAGC 2700
QY 2932 TTGTTACAAAGACGGGACGCTGTTCCGTACGACATGGAGGTGCCAGACTCGGGCATC 2991
Db 2701 TTGTTACAAAGACGGGACGCTGTTCCGTACGACATGGAGGTGCCAGACTCGGGCATC 2760
QY 2992 GACCTGAGTGCACACTGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTTCAAGCAT 3051
Db 2761 GACCTGAGTGCACACTGGCCCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTTCAAGCAT 2820
```

```
QY 3052 GGCCAGCTGGAGAAACAGGCCCTTAACCTTGCCCTCCACCGCCGGCTCCACACTGCGCGAAA 3111
Db 2821 GGCCAGCTGGAGAAACAGGCCCTTAACCTTGCCCTCCACCGCCGGCTCCACACTGCGCGAAA 2880
QY 3112 GCAGCCTTCTGCTCGGTGACAGATGCTGCCCCTGAACACACAGGCTCAGCGTTCCCAAG 3171
Db 2881 GCAGCCTTCTGCTCGGTGACAGATGCTGCCCCTGAACACACAGGCTCAGCGTTCCCAAG 2940
QY 3172 GGATTGCGAGCCCCCGGCTCACAGTGGAAACCAAGGGCTCGCAGCAGCAAGGTGGGGCC 3231
Db 2941 GGATTGCGAGCCCCCGGCTCACAGTGGAAACCAAGGGCTCGCAGCAGCAAGGTGGGGCC 3000
QY 3232 AAGCAGAAATGCTCCCAAGATTTACACTGAGCCCTGCCCCACCTCTCTGTAACCAAT 3291
Db 3001 AAGCAGAAATGCTCCCAAGATTTACACTGAGCCCTGCCCCACCTCTCTGTAACCAAT 3060
QY 3292 CCGCAGCTGAAGAGACAGAGGAGGATGGCAGAGTTACCTGGGGAACAAACAGGGA 3351
Db 3061 CCGCAGCTGAAGAGACAGAGGAGGATGGCAGAGTTACCTGGGGAACAAACAGGGA 3120
QY 3352 TCTTTTCTGCCCCCTGCTCCAGTCGAGTTGGCCTGA 3387
Db 3121 TCTTTTCTGCCCCCTGCTCCAGTCGAGTTGGCCTGA 3156

RESULT 9
AR068515 LOCUS AR068515 3156 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5854003.
ACCESSION AR068515
VERSION AR068515.1 GI:6000722
KEYWORDS SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3156)
AUTHORS Rothe,M. and Wu,L.
TITLE Screening method for agents that modulate human NIK activity
JOURNAL Patent: US 5854003-A 1 29-DEC-1998;
FEATURES Location/Qualifiers
source 1..3156
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 68.5%; Score 3149.6; DB 6; Length 3156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 232 ATGCAGTGTATGGAATGCGCTGCCAGGTGCCCTCGCTCAGCAGTGGGGCAGCAGAAAG 291
Db 1 ATGCAGTGTATGGAATGCGCTGCCAGGTGCCCTCGCTCAGCAGTGGGGCAGCAGAAAG 60
QY 292 GAATCTCCCAAGCAAGAGGAGAGACCCCGCACTGGGGGAAGAAACAGAGTCCGTCTTAC 351
Db 61 GAATCTCCCAAGCAAGAGGAGAGACCCCGCACTGGGGGAAGAAACAGAGTCCGTCTTAC 120
QY 352 AAGCTTGAGCCGTGGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 411
Db 121 AAGCTTGAGCCGTGGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 180
QY 412 GTGATTACCAAGGCGCAGCAAGGAAGGCTCCAGAGCAGGGCCAGCTGCCATCTCTATC 471
Db 181 GTGATTACCAAGGCGCAGCAAGGAAGGCTCCAGAGCAGGGCCAGCTGCCATCTCTATC 240
QY 472 ATCGCCAGGCTGAGTGTGAGAAATAGCAAGATTACGCCCCACCTTTTCAGAACGCAAT 531
Db 241 ATCGCCAGGCTGAGTGTGAGAAATAGCAAGATTACGCCCCACCTTTTCAGAACGCAAT 300
QY 532 TTCAATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAT 591
Db 301 TTCAATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAT 360
```

QY 592 GTGGCCCATCTACAGAGCGCAAAATGGCCCGTGTGTGTGGAGGGGAAAGCGTCGCAGC 651  
DB |||||  
QY 361 GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAGGGGAAAGCGTCGCAGC 420  
DB |||||  
QY 652 AAAGCCCGGAAGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAAGGAGTGCC 711  
DB |||||  
QY 421 AAAGCCCGGAAGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCAAGGAGTGCC 480  
DB |||||  
QY 712 TTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGAGGAG 771  
DB |||||  
QY 481 TTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGAGGAG 540  
DB |||||  
QY 772 GATGAGTCTCCACTCGGGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTG 831  
DB |||||  
QY 541 GATGAGTCTCCACTCGGGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCTCTG 600  
DB |||||  
QY 832 AAGGAACCAAGGCTTTGGGCAACTCTGTTTAAAGCAGTTGGCGAGGGCTACGGCCGGCT 891  
DB |||||  
QY 601 AAGGAACCAAGGCTTTGGGCAACTCTGTTTAAAGCAGTTGGCGAGGGCTACGGCCGGCT 660  
DB |||||  
QY 892 CTGCTCGATCAGAACTCCACAACTGATCAGCCCTTGCATGCTGAAACCAAGTGAG 951  
DB |||||  
QY 661 CTGCTCGATCAGAACTCCACAACTGATCAGCCCTTGCATGCTGAAACCAAGTGAG 720  
DB |||||  
QY 952 AAAGTGCACACCCCGAGGAGCGGCGCCCTGCCCCCTGCCACGACCCCTTCCCTTAT 1011  
DB |||||  
QY 721 AAAGTGCACACCCCGAGGAGCGGCGCCCTGCCCCCTGCCACGACCCCTTCCCTTAT 780  
DB |||||  
QY 1012 AGCAGACTGCCTCATCCCTTCCCATTTCCATCCACCTCTCCAGCCCTGGAACCTCACCTCTG 1071  
DB |||||  
QY 781 AGCAGACTGCCTCATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCTCTG 840  
DB |||||  
QY 1072 GAGTCTCTCTGGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACAC 1131  
DB |||||  
QY 841 GAGTCTCTCTGGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACAC 900  
DB |||||  
QY 1132 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGGCC 1191  
DB |||||  
QY 901 CTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGGCC 960  
DB |||||  
QY 1192 AGCTGCCTGTCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGATGCT 1251  
DB |||||  
QY 961 AGCTGCCTGTCTCGTGGTCCCATGAGAAATTTCTGTGGAGGAATACCTAGTGATGCT 1020  
DB |||||  
QY 1252 CTGCAAGGAGGCTGAGCTCAAGCCAGGCGCCAGAGCTTGCCTGAGGAGCTGCGCAAGCCTGG 1311  
DB |||||  
QY 1021 CTGCAAGGAGGCTGAGCTCAAGCCAGGCGCCAGAGCTTGCCTGAGGAGCTGCGCAAGCCTGG 1080  
DB |||||  
QY 1312 GCAGACAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGCTGTC 1371  
DB |||||  
QY 1081 GCAGACAGGGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGCTGTC 1140  
DB |||||  
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATATGAGTACCGAGAGAAAGTCCACTGGGCC 1431  
DB |||||  
QY 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATATGAGTACCGAGAGAAAGTCCACTGGGCC 1200  
DB |||||  
QY 1432 AGCCACAGCTCGGCTGGGAGAGGCTCCCTTGGAGAGTGCACAGGATGAGGACAG 1491  
DB |||||  
QY 1201 AGCCACAGCTCGGCTGGGAGAGGCTCCCTTGGAGAGGTCACAGGATGAGGACAG 1260  
DB |||||  
QY 1492 CAGACTGGCTTCAGTGGCTGTCAAAAAGTGGCGCTGGAAAGTATTTCCGGCAGAGGAG 1551  
DB |||||  
QY 1261 CAGACTGGCTTCAGTGGCTGTCAAAAAGTGGCGCTGGAAAGTATTTCCGGCAGAGGAG 1320  
DB |||||  
QY 1552 CTGATGGCATGTGCAGGATGACCTCAACAGAAATTTGCTCCCTTGTATGGAGCTGTGAGA 1611  
DB |||||  
QY 1321 CTGATGGCATGTGCAGGATGACCTCAACAGAAATTTGCTCCCTTGTATGGAGCTGTGAGA 1380  
DB |||||  
QY 1612 GAAGGCCCTTGGGTCAACATCTTCATGAGAGCTCTGGAAGTGGCTCCCTGGGCCAGCTG 1671  
DB |||||  
QY 1381 GAAGGCCCTTGGGTCAACATCTTCATGAGAGCTCTGGAAGTGGCTCCCTGGGCCAGCTG 1440  
DB |||||  
QY 1672 GTCAAGGAGCAGGGCTGTCTCCAGAGGAGCCGGGCCCTGTACTACCTGGGCCAGGCCCTG 1731

DB |||||  
QY 1441 GTCAAGGAGCAGGGCTGTCTCCAGAGAGCCGGGCCCTGTACTACCTGGGCCAGGCCCTG 1500  
DB |||||  
QY 1732 GAGGGTCTGGAATACCTCCACTCAGAAAGGATCTGTCATGGGGACGTCAAAGCTGACAAAC 1791  
DB |||||  
QY 1501 GAGGGTCTGGAATACCTCCACTCAGAAAGGATCTGTCATGGGGACGTCAAAGCTGACAAAC 1560  
DB |||||  
QY 1792 GTGCTCTCTGTCAGCGATGGGAGCAGCGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851  
DB |||||  
QY 1561 GTGCTCTCTGTCAGCGATGGGAGCAGCGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620  
DB |||||  
QY 1852 CTTCAACTGATGCTGGGAAAGTCTCTGTCACAGGGACTACATCCCTGGGCACAGAG 1911  
DB |||||  
QY 1621 CTTCAACTGATGCTGGGAAAGTCTCTGTCACAGGGACTACATCCCTGGGCACAGAG 1680  
DB |||||  
QY 1912 ACCCACATGCTCGGAGGTGCTGCGGAGGAGCTGCGAGCGCAAGGTGATGTCGTGG 1971  
DB |||||  
QY 1681 ACCCACATGCTCGGAGGTGCTGCGGAGGAGCTGCGAGCGCAAGGTGATGTCGTGG 1740  
DB |||||  
QY 1972 AGCAGCTGCTATGATGCTGCATGCTCAACCGCTGSCCACCCTGTGACTCAGTTCTTC 2031  
DB |||||  
QY 1741 AGCAGCTGCTATGATGCTGCATGCTCAACCGCTGSCCACCCTGTGACTCAGTTCTTC 1800  
DB |||||  
QY 2032 CGAGGGCGCTCTGCTCAAGATTGCCAGCGAGCCCTCGCCTGTGAGGGAGATCCCAACC 2091  
DB |||||  
QY 1801 CGAGGGCGCTCTGCTCAAGATTGCCAGCGAGCCCTCGCCTGTGAGGGAGATCCCAACC 1860  
DB |||||  
QY 2092 TCCTGCGCCCTCTCAGAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCCATCCAC 2151  
DB |||||  
QY 1861 TCCTGCGCCCTCTCAGAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCCATCCAC 1920  
DB |||||  
QY 2152 CGCTGTCTGAGCGGAGCTGGGAGGAGGAGTGAACCGGGCACTACAGCAAGTGGAGGT 2211  
DB |||||  
QY 1921 CGCTGTCTGAGCGGAGCTGGGAGGAGGAGTGAACCGGGCACTACAGCAAGTGGAGGT 1980  
DB |||||  
QY 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCCCAAAATCAAGCC 2271  
DB |||||  
QY 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCCCAAAATCAAGCC 2040  
DB |||||  
QY 2272 AATTACACACAGACCTCTCATGCCAGCGGAGAGAGCTTCGCCAAGGGGCCCGAGGGCCC 2331  
DB |||||  
QY 2041 AATTACACACAGACCTCTCATGCCAGCGGAGAGAGCTTCGCCAAGGGGCCCGAGGGCCC 2100  
DB |||||  
QY 2332 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAG 2391  
DB |||||  
QY 2101 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACACAGAG 2160  
DB |||||  
QY 2392 CCCCAGAGCCAAACAAAGTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2451  
DB |||||  
QY 2161 CCCCAGAGCCAAACAAAGTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2220  
DB |||||  
QY 2452 GAAACCTTTACCTCTGTCTCTCTCTGAGCGAGCCCTGCGAGAAACCCAGAGCTCAGCAGAG 2511  
DB |||||  
QY 2221 GAAACCTTTACCTCTGTCTCTCTCTGAGCGAGCCCTGCGAGAAACCCAGAGCTCAGCAGAG 2280  
DB |||||  
QY 2512 CGGAAAGCAACCGTCCCGAGCAGGAACTGAGCAGAGCTGGAATAGAAATTTCTCTCAAC 2571  
DB |||||  
QY 2281 CGGAAAGCAACCGTCCCGAGCAGGAACTGAGCAGAGCTGGAATAGAAATTTCTCTCAAC 2340  
DB |||||  
QY 2572 AGCTGTCTCCAGCCATTTTCTGTGAGGAGCAGGAGCAAAATTTCTCTGCTCTGAGCATC 2631  
DB |||||  
QY 2341 AGCTGTCTCCAGCCATTTTCTGTGAGGAGCAGGAGCAAAATTTCTCTGCTCTGAGCATC 2400  
DB |||||  
QY 2632 GACAGCCTCTCCCTGTGCGATGACAGTGAAGAACCCATCAAGGCTCTCTCAAGCTCG 2691  
DB |||||  
QY 2401 GACAGCCTCTCCCTGTGCGATGACAGTGAAGAACCCATCAAGGCTCTCTCAAGCTCG 2460  
DB |||||  
QY 2692 CGGAGACACCTGAGCTCAGGGGTACACTCTGGAGCAGCAGGCGGAGCTCGAAGCTCC 2751  
DB |||||  
QY 2461 CGGAGACACCTGAGCTCAGGGGTACACTCTGGAGCAGCAGGCGGAGCTCGAAGCTCC 2520  
DB |||||  
QY 2752 AGCTGGAACATGCTGCTGCGCCCGGGGCGGCCACCGACACCCCAAGCTATTTTCAATGCT 2811



Db 2521 AGCTGGAACATGGTCTGGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580  
Qy 2812 GTGAAAGTCCAAATACAGTCTCTTAATAGTGAACACCTGCACATCCCGGAGTTCCACCGG 2871  
Db 2581 GTGAAAGTCCAAATACAGTCTCTTAATAGTGAACACCTGCACATCCCGGAGTTCCACCGG 2640  
Qy 2872 GTCAAAGTGGGAGACATCCCACTGGGATCAGCAGCAGATCCCAAGTGCAGCTTCCAGC 2931  
Db 2641 GTCAAAGTGGGAGACATCCCACTGGGATCAGCAGCAGATCCCAAGTGCAGCTTCCAGC 2700  
Qy 2932 TTGGTACCAAAAGACGGCAGCCTGTTCGTACGACATGAGAGTGCACAGCTCGGGCATC 2991  
Db 2701 TTGGTACCAAAAGACGGCAGCCTGTTCGTACGACATGAGAGTGCACAGCTCGGGCATC 2760  
Qy 2992 GACCTGCAGTGCACATGCGCCCTCTGATGGCAGCTTTCGCTGGAGTGGAGGGTCAAGCAT 3051  
Db 2761 GACCTGCAGTGCACATGCGCCCTCTGATGGCAGCTTTCGCTGGAGTGGAGGGTCAAGCAT 2820  
Qy 3052 GGCCAGCTGGAGAACAGSCCTTAACCTTGCCTTCACCGCGGCTCCACATCGCGGNA 3111  
Db 2821 GGCCAGCTGGAGAACAGSCCTTAACCTTGCCTTCACCGCGGCTCCACATCGCGGNA 2880  
Qy 3112 GCAGCTTCTCTGCTCGGTGCAGATGCTGCTGAAACACAGGCTCAGCCGTTCCCGAGG 3171  
Db 2881 GCAGCTTCTCTGCTCGGTGCAGATGCTGCTGAAACACAGGCTCAGCCGTTCCCGAGG 2940  
Qy 3172 GGATTGCCAGCCCCCGCTCACAGTGGGAACAGGGCTCGCAGCAAGTGGGGGC 3231  
Db 2941 GGATTGCCAGCCCCCGCTCACAGTGGGAACAGGGCTCGCAGCAAGTGGGGGC 3000  
Qy 3232 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGAAAAACAT 3291  
Db 3001 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGAAAAACAT 3060  
Qy 3292 CCGCCAGCTGAAGACAGAGAGGAGTGGCAGGATTTACCTGGGAAACAAACAGGA 3351  
Db 3061 CCGCCAGCTGAAGACAGAGAGGAGTGGCAGGATTTACCTGGGAAACAAACAGGA 3120  
Qy 3352 TCTTTTCTGCCCCCTGCTCCAGTCCAGTGGCTCGA 3387  
Db 3121 TCTTTTCTGCCCCCTGCTCCAGTCCAGTGGCTCGA 3156

RESULT 10  
LOCUS BD062401 3156 bp DNA linear PAT 27-AUG-2002  
DEFINITION NIK proteins, nucleic acids and methods.  
ACCESSION BD062401  
VERSION BD062401.1 GI:22608004  
KEYWORDS JP 2001510348-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Rothe.M. and Wu.L.  
TITLE NIK proteins, nucleic acids and methods  
JOURNAL Patent: JP 2001510348-A 1 31-JUL-2001;  
TULARIK INC  
COMMENT PN JP 2001510348-A/1  
PD 31-JUL-2001  
PF 02-JUL-1998 JP 1999507409  
PR 03-JUL-1997 US 08/887518  
PI MIKE ROTHE, LIN WU  
PC  
C07K7/06, C07K7/08, C07K14/47, C12N1/00, C12N5/10, C12N15/12, C12N15/ PC  
63,  
PC C12P21/02, G01N33/53  
CC Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
1. .3156

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 68.5%; Score 3149.6; DB 6; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 232 ATGCGAGTGTGAAATGGCTGCGCAGGTGCGCCCTGGCTCAGCAGTGGGCGAGCAGAG 291  
Db 1 ATGCGAGTGTGAAATGGCTGCGCAGGTGCGCCCTGGCTCAGCAGTGGGCGAGCAGAG 60  
Qy 292 GAATCCCAAGCCAAAGGAGAGACGCCGCCACTTGGGGAGAAACAGAGTCCGTCTAC 351  
Db 61 GAATCCCAAGCCAAAGGAGAGACGCCGCCACTTGGGGAGAAACAGAGTCCGTCTAC 120  
Qy 352 AAGCTTGAGCGCGTGGAGAGAGCCCTGTGTTCTCGCGAAGTGGAGATCCTGAATGAC 411  
Db 121 AAGCTTGAGCGCGTGGAGAGAGCCCTGTGTTCTCGCGAAGTGGAGATCCTGAATGAC 180  
Qy 412 GTGATTACCAAGGCGCAGCAAGAAAGGCTCCGAGCGAGGCGCAGCTGCCATCTCTATC 471  
Db 181 GTGATTACCAAGGCGCAGCAAGAAAGGCTCCGAGCGAGGCGCAGCTGCCATCTCTATC 240  
Qy 472 ATGCCCGAGGCTGAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTCAGAAAGCAT 531  
Db 241 ATGCCCGAGGCTGAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTCAGAAAGCAT 300  
Qy 532 TTGATCCCTGGTCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCCAACAT 591  
Db 301 TTGATCCCTGGTCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCCAACAT 360  
Qy 592 GTGCGCCATGCTACAGAGGGCAAAATGCGCCGTGTGTTGGAAGGAAACGCTCGCAGC 651  
Db 361 GTGCGCCATGCTACAGAGGGCAAAATGCGCCGTGTGTTGGAAGGAAACGCTCGCAGC 420  
Qy 652 AAAGCCCGAAGAAACGGAAGAGAGTCAAGTCCCTGGCTCATGCAAGGAGTGGCC 711  
Db 421 AAAGCCCGAAGAAACGGAAGAGAGTCAAGTCCCTGGCTCATGCAAGGAGTGGCC 480  
Qy 712 TTGSCCAAAACCCCTCCCGAGGACCCCTGAGCAGAGAGTGCACATCCAGTGGAGAG 771  
Db 481 TTGSCCAAAACCCCTCCCGAGGACCCCTGAGCAGAGAGTGCACATCCAGTGGAGAG 540  
Qy 772 GATGAGTCTCACCTCGCGGCCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTG 831  
Db 541 GATGAGTCTCACCTCGCGGCCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTG 600  
Qy 832 AAGGAACCCAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCCGGCT 891  
Db 601 AAGGAACCCAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCCGGCT 660  
Qy 892 CTGCTCGATCAGAACTCCCAAACTGATCAGCCCTTGCATCTGAAACACAGTGTGG 951  
Db 661 CTGCTCGATCAGAACTCCCAAACTGATCAGCCCTTGCATCTGTAACACAGTGTGG 720  
Qy 952 AAAGTCACACCCCGCAGGAGGAGGCGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTCAT 1011  
Db 721 AAAGTCACACCCCGCAGGAGGAGGCGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTCAT 780  
Qy 1012 AGCAGACTGCTCATCTCCCTTCCCACTTCCAGCCCTTCCAGCCCTTCCAGCCCTCTG 1071  
Db 781 AGCAGACTGCTCATCTCCCTTCCCACTTCCAGCCCTTCCAGCCCTTCCAGCCCTCTG 840  
Qy 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCAGCAAGAAACCTTGGCTTGACCCACAC 1131  
Db 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCAGCAAGAAACCTTGGCTTGACCCACAC 900  
Qy 1132 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGGCGCCCTGCTGCGCCACACACCTGGAGGCC 1191  
Db 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGGCGCCCTGCTGCGCCACACACCTGGAGGCC 960

QY 1192 AGCTGCCTGTCTCGTGGTGCCCATGAGAAAGTTTTCGTGGAGGAATACCTAGTGCATGCT 1251  
DB 961 AGCTGCCCTGTCTCGTGGTGCCCATGAGAAAGTTTTCGTGGAGGAATACCTAGTGCATGCT 1020  
QY 1252 CTGCAAGGAGCGGTGAGTCTAAGCCAGGCCCAACAGCCCTGACAGCCTGGCCAAAGACCTGG 1311  
DB 1021 CTGCAAGGAGCGGTGAGTCTAAGCCAGGCCCAACAGCCCTGACAGCCTGGCCAAAGACCTGG 1080  
QY 1312 GCAGCAGCGGGCTCAGATCCCGGGAGCCAGCCCCCAAAATCTGAGGACAACGAGGGGTGC 1371  
DB 1081 GCAGCAGCGGGCTCAGATCCCGGGAGCCAGCCCCCAAAATCTGAGGACAACGAGGGGTGC 1140  
QY 1372 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1431  
DB 1141 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1200  
QY 1432 ACGCAACCAAGCTCCGCTGGGCAGAGGCTCTTCGAGAGGTGCACAGATGAGAGCAAG 1491  
DB 1201 ACGCAACCAAGCTCCGCTGGGCAGAGGCTCTTCGAGAGGTGCACAGATGAGAGCAAG 1260  
QY 1492 CAGACTGGCTTCAGTGGCGCTGTCAAAAGGTGCGCTGGAGTATTTGGGCGAGAGG 1551  
DB 1261 CAGACTGGCTTCAGTGGCGCTGTCAAAAGGTGCGGCTGGAAGTATTTGGGCGAGAGG 1320  
QY 1552 CTGATGGCATGTGCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGA 1611  
DB 1321 CTGATGGCATGTGCAGGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGA 1380  
QY 1612 GAAGGGCTTGGGTCAACATCTTCATGAGCTGTGGAAGTGGCTTCCTTGGGCCAGCTG 1671  
DB 1381 GAAGGGCTTGGGTCAACATCTTCATGAGCTGTGGAAGTGGCTTCCTTGGGCCAGCTG 1440  
QY 1672 GTCAAGGACAGGGCTGTCTCCAGAGGACCGGCGCTGTACTACTGGGCCAGGCCCTG 1731  
DB 1441 GTCAAGGACAGGGCTGTCTCCAGAGGACCGGCGCTGTACTACTGGGCCAGGCCCTG 1500  
QY 1732 GAGGGTCTGGAATACCTCCACTCACGAAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAA 1791  
DB 1501 GAGGGTCTGGAATACCTCCACTCACGAAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAA 1560  
QY 1792 GTGCTCTCTCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCATGCTGTGTGT 1851  
DB 1561 GTGCTCTCTCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCATGCTGTGTGT 1620  
QY 1852 CTTCAACCTGATGGCTGGGAAGTCTTGTCTCACAGGGAGCTACATCCCTGGCAGAG 1911  
DB 1621 CTTCAACCTGATGGCTGGGAAGTCTTGTCTCACAGGGAGCTACATCCCTGGCAGAG 1680  
QY 1912 ACCCAGATGGCTCCGAGGTGGTGTGGCAGAGCTGCGAGCGCCAAAGGTGGATGTCTGG 1971  
DB 1681 ACCCAGATGGCTCCGAGGTGGTGTGGCAGAGCTGCGAGCGCCAAAGGTGGATGTCTGG 1740  
QY 1972 AGCAGTGTGTATGATGTGCATGTCTCAACGGCTGCGACCCCTGGACTCAGTTCTTTC 2031  
DB 1741 AGCAGTGTGTATGATGTGCATGTCTCAACGGCTGCGACCCCTGGACTCAGTTCTTTC 1800  
QY 2032 CGAGGGCGCTCTGCTCTCAAGATTGCGAGGAGCTCCGCTCTGAGGGAGATCCACCC 2091  
DB 1801 CGAGGGCGCTCTGCTCTCAAGATTGCGAGGAGCTCCGCTCTGAGGGAGATCCACCC 1860  
QY 2092 TCTGCGCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCAC 2151  
DB 1861 TCTGCGCCCTCTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCAC 1920  
QY 2152 CGCGTGTGTGACGGAGCTGGAGGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGT 2211  
DB 1921 CGCGTGTGTGACGGAGCTGGAGGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGT 1980  
QY 2212 CTGAAGGCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCCAATCAAGCC 2271  
DB 1981 CTGAAGGCCCTTGGAGGGAGATATAAGAACCAAGACATCCACCGCCCAATCAAGCC 2040  
QY 2272 AATTACCAACGACCCCTCCATGTCACCGCAGAGAGGCTTTTCGCAAGGGCCCCAGGGCCCC 2331

DB 2041 AATTACCAACGACCCCTCCATGCCACCGAGAGAGCTTTTCGCAAGGGCCCCAGGGCCCC 2100  
QY 2332 CGGCCAGCTGAGGAGACAACAGCAGAGCCCTTAAGCTCTCAGCCTCTCTCTCCCAACAGAG 2391  
DB 2101 CGGCCAGCTGAGGAGACAACAGCAGAGCCCTTAAGCTCTCAGCCTCTCTCTCCCAACAGAG 2160  
QY 2392 CCCCAGAGCCAAACAAGTCTCTCTCCCTTGAATTGAGCAAGGAGGAGTCTGGGATGTGG 2451  
DB 2161 CCCCAGAGCCAAACAAGTCTCTCTCCCTTGAATTGAGCAAGGAGGAGTCTGGGATGTGG 2220  
QY 2452 GAAACCTTACTCTGTCTCTCTGGAGCAGCCCTTCGAGAAACCCAGCTCACAGAG 2511  
DB 2221 GAAACCTTACTCTGTCTCTCTGGAGCAGCCCTTCGAGAAACCCAGCTCACAGAG 2280  
QY 2512 CGGAAAGCAACCGTCCCGGAGCAGAACTGCAGCAGCTGGAAATAGAAATTAATTCCTCAAC 2571  
DB 2281 CGGAAAGCAACCGTCCCGGAGCAGAACTGCAGCAGCTGGAAATAGAAATTAATTCCTCAAC 2340  
QY 2572 AGCCTGTCCCAGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTGCTCAGCATC 2631  
DB 2341 AGCCTGTCCCAGCCATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTGCTCAGCATC 2400  
QY 2632 GACAGCCTCTCCCTGTGCGATGACAGTGAAGAAACCAATCAAGGCTCTCTCAAGCTCG 2691  
DB 2401 GACAGCCTCTCCCTGTGCGATGACAGTGAAGAAACCAATCAAGGCTCTCTCAAGCTCG 2460  
QY 2692 CGGCAACCCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCGCCGAGGCTCGAAGCTCC 2751  
DB 2461 CGGCAACCCCTGAGCTCAGGCGTACACTCTCTGGAGCAGCAGCGCCGAGGCTCGAAGCTCC 2520  
QY 2752 AGCTGGAACTGTGTGTCGCGCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2811  
DB 2521 AGCTGGAACTGTGTGTCGCGCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580  
QY 2812 GTGAAAGTCCAAATACAGTCTCTTAATGGTGAACAACCTGCAATCCGCGAGTTCACCGG 2871  
DB 2581 GTGAAAGTCCAAATACAGTCTCTTAATGGTGAACAACCTGCAATCCGCGAGTTCACCGG 2640  
QY 2872 GTCAAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCAGC 2931  
DB 2641 GTCAAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCGAGCTTCAGC 2700  
QY 2932 TTGGTCAACAAAGAGCGGCGAGCTGTTGCTACGACATGGAGGTGCCAGATCTCGGGCATC 2991  
DB 2701 TTGGTCAACAAAGAGCGGCGAGCTGTTGCTACGACATGGAGGTGCCAGATCTCGGGCATC 2760  
QY 2992 GACCTGCAAGTGCACATCGGCCCTGTATGGCAGCTTCGCTGAGCTGAGGCTCAAGCAT 3051  
DB 2761 GACCTGCAAGTGCACATCGGCCCTGTATGGCAGCTTCGCTGAGCTGAGGAGTCAAGCAT 2820  
QY 3052 GGCAGCTGGAGAAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCCCGGAAA 3111  
DB 2821 GGCAGCTGGAGAAACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCCCGGAAA 2880  
QY 3112 GCAGCCTTCTGCTCGGTGCAGATGCTGCTGAAAAACAGGCTCAGCGTTCGCCAGG 3171  
DB 2881 GCAGCCTTCTGCTCGGTGCAGATGCTGCTGAAAAACAGGCTCAGCGTTCGCCAGG 2940  
QY 3172 GGAATGCGAGCCCCCGGCTCAAGTGGGAACAGGGCTTCGACAGCAGAGGTGGGGC 3231  
DB 2941 GGAATGCGAGCCCCCGGCTCAAGTGGGAACAGGGCTTCGACAGCAGAGGTGGGGC 3000  
QY 3232 AAGCAGAAATGCTCCAGGATTTACACTGAGCCCTGCCACCCCTGCTGAAAAAACAAT 3291  
DB 3001 AAGCAGAAATGCTCCAGGATTTACACTGAGCCCTGCCACCCCTGCTGAAAAAACAAT 3060  
QY 3292 CGGCCACCTGAGAGACAGAAAGGAGTGGCAGGAGTTACTTGGGAAAACAAAAACAGGA 3351  
DB 3061 CGGCCACCTGAGAGACAGAAAGGAGTGGCAGGAGTTACTTGGGAAAACAAAAACAGGA 3120  
QY 3352 TCTTTTCTGCCCCCTGCTCCAGTCCAGTTGGCTGA 3387

Db	3121	TCTTTTTCGCCCTGCTCCAGTCGAGTGCGCCTGA	3156
RESULT 11	AR429673		
LOCUS	AR429673	2844 bp	DNA linear PAT 18-DEC-2003
DEFINITION	Sequence 2 from patent US 6645728.		
ACCESSION	AR429673		
VERSION	AR429673.1	GI:40189976	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 2844)		
TITLE	Greene,W.C., Lin,X. and Gelezuinas,R. Inhibitor of the inflammatory response induced by TNF.alpha. and IL-1		
JOURNAL	Patent: US 6645728-A 2 11-NOV-2003;		
FEATURES	Location/Qualifiers		
source	1..2844		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	61.7%; Score 2837.6; DB 6; Length 2844;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 2840; Conservative	0; Mismatches 4; Indels 0; Gaps 0;		
Qy	232	ATGGCAGTGATGGAATGGCTGCCCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGNAG	291
Db	1	ATGGCAGTGATGGAATGGCTGCCCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGNAG	60
Qy	292	GAACTCCCCAAGCCAAAGGAGAAGACCAGCCGCACCTGGGGGAAGAAAACAGAGCTCCGCTCTAC	351
Db	61	GAACTCCCCAAGCCAAAGGAGAAGACCAGCCGCACCTGGGGGAAGAAAACAGAGCTCCGCTCTAC	120
Qy	352	AAGCTTGAGGCGGTGGAGAAAGAGCCCTGTGTCTGCGGAAAGTGGGAGATCCTGGAATGAC	411
Db	121	AAGCTTGAGGCGGTGGAGAAAGAGCCCTGTGTCTGCGGAAAGTGGGAGATCCTGGAATGAC	180
Qy	412	GTGATTACCAAGGCGACAGCCNAGGAGGCTCCGAGGACAGGCGCAGCTGCCATCTCTATC	471
Db	181	GTGATTACCAAGGCGCACAGCCNAGGAGGCTCCGAGGACAGGCGCAGCTGCCATCTCTATC	240
Qy	472	ATCGCCCGAGGCTGAGTGTGAGAAATAGCCAAGAGTTTACGCCCCACCTTTTTCAGAAACGCATT	531
Db	241	ATCGCCCGAGGCTGAGTGTGAGAAATAGCCAAGAGTTTACGCCCCACCTTTTTCAGAAACGCATT	300
Qy	532	TTCATCGCTGGGTCAAACAAGTAGTACAGCCAGTCCGAGAGTCTTGATCAGATFCCCACAACAT	591
Db	301	TTCATCGCTGGGTCAAACAAGTAGTACAGCCAGTCCGAGAGTCTTGATCAGATFCCCACAACAT	360
Qy	592	GTGGCCCATGCTACAGAGGGCAAAATGGCCCGGTGTGTGTTGGAAGGAAAGCGTCGAGC	651
Db	361	GTGGCCCATGCTACAGAGGGCAAAATGGCCCGGTGTGTGTTGGAAGGAAAGCGTCGAGC	420
Qy	652	AAAGCCCGGAAGAAACGGAAGAACAGAGCTCAAAGTCCCTGGCTCATGCAGAGATGGCC	711
Db	421	AAAGCCCGGAAGAAACGGAAGAACAGAGCTCAAAGTCCCTGGCTCATGCAGAGATGGCC	480
Qy	712	TTGGCCAAACCCCTCCCCAGGACCCCTTGAGCAGGAGAGCTGCACCATTCCAGTGCAGGAG	771
Db	481	TTGGCCAAACCCCTCCCCAGGACCCCTTGAGCAGGAGAGCTGCACCATTCCAGTGCAGGAG	540
Qy	772	GATGAGTCTGACTCGGCGGCCCONATGTTAGAAAAACACCCCGAGTTCACCAAGCCTCTG	831
Db	541	GATGAGTCTTCACTCGGCGGCCCATATGTTAGAAAAACACCCCGAGTTCACCAAGCCTCTG	600
Qy	832	AAGGAACACAGGCTTGGGCAACTCTGTTTTTAAGCAGCTTGGCGAGGSCCTACGGCCGGCT	891
Db	601	AAGGAACACAGGCTTGGGCAACTCTGTTTTTAAGCAGCTTGGCGAGGSCCTACGGCCGGCT	660
Qy	892	CTGCGCTCGATCAGAACTCCCAAACTGATCAGGCCCTTTGCAATGTCTGAAACACAGTGTGG	951

Db 1741 AGCAGCTGCTGATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGCACTCAGTTCTTC 1800  
QY 2032 CGAGGGCCGCTCTGCTCAAGATTGCCAGCGAGCTTCGCCCTGTGAGGGAGATCCACCC 2091  
Db 1801 CGAGGGCCGCTCTGCTCAAGATTGCCAGCGAGCTTCGCCCTGTGAGGGAGATCCACCC 1860  
QY 2092 TCTGTGCGCCCTCTCACAGCCCGAGCCATTCGAAGAGGGCTGAGAAAGAGCCCATCCAC 2151  
Db 1861 TCTGTGCGCCCTCTCACAGCCCGAGCCATTCGAAGAGGGCTGAGAAAGAGCCCATCCAC 1920  
QY 2152 CGCTGTCTGACGCGAGCTGGAGGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGT 2211  
Db 1921 CGCTGTCTGACGCGAGCTGGAGGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGT 1980  
QY 2212 CTGAAGAGCCCTTGGAGGGGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCC 2271  
Db 1981 CTGAAGAGCCCTTGGAGGGGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCC 2040  
QY 2272 AATTACCACAGACCTTCATGCCAGCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCC 2331  
Db 2041 AATTACCACAGACCTTCATGCCAGCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCC 2100  
QY 2332 CGGCAGCTGAGAGACAAACAGGAGAGCCCTTAAGCTTCAGCCTCTCTTCCCAACAGAG 2391  
Db 2101 CGGCAGCTGAGAGACAAACAGGAGAGCCCTTAAGCTTCAGCCTCTCTTCCCAACAGAG 2160  
QY 2392 CCCCGAGGCCAAACAAAGTCTCCTCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGG 2451  
Db 2161 CCCCGAGGCCAAACAAAGTCTCCTCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGG 2220  
QY 2452 GAACCCCTTACCTCTGTCTCCTCGAGCGAGCCCTTGCCAGAAACCCAGCTCACCAGAG 2511  
Db 2221 GAACCCCTTACCTCTGTCTCCTCGAGCGAGCCCTTGCCAGAAACCCAGCTCACCAGAG 2280  
QY 2512 CGGAAGACACCGTCCCGAGCAGGAACTGCAGAGCTGGAAATAGAATATTCTCTCAAC 2571  
Db 2281 CGGAAGACACCGTCCCGAGCAGGAACTGCAGAGCTGGAAATAGAATATTCTCTCAAC 2340  
QY 2572 AGCTGTCTCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATCTCTCGCTCAGCATC 2631  
Db 2341 AGCTGTCTCCAGCAATTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATC 2400  
QY 2632 GACAGCCTCTCCTGTCTGAGTGAACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCG 2691  
Db 2401 GACAGCCTCTCCTGTCTGAGTGAACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCG 2460  
QY 2692 CGGGACACCTGAGCTCAGGCGTACATCTCTGAGAGCAGCGCGAGGCTCGAAGCTCC 2751  
Db 2461 CGGGACACCTGAGCTCAGGCGTACATCTCTGAGAGCAGCGCGAGGCTCGAAGCTCC 2520  
QY 2752 AGCTGGAACATGCTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTCAATGGT 2811  
Db 2521 AGCTGGAACATGCTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTCAATGGT 2580  
QY 2812 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGGAGTTTCCACCGG 2871  
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGGAGTTTCCACCGG 2640  
QY 2872 GTCAAAGTGGAGACATCGGCATCAGCAGCGCAGATCCCAAGCTTCAGC 2931  
Db 2641 GTCAAAGTGGAGACATCGGCATCAGCAGCGCAGATCCCAAGCTTCAGC 2700  
QY 2932 TTGGTCAACAAAGACGGCAGCCTCTGCTACGACATGGAGTCCAGACTCGGGCATC 2991  
Db 2701 TTGGTCAACAAAGACGGCAGCCTCTGCTACGACATGGAGTCCAGACTCGGGCATC 2760  
QY 2992 GACCTGCAGTGACACTGCGCCCTGATGGCAGCTTCGCTTGAGCTGGAGGTCAGCAT 3051  
Db 2761 GACCTGCAGTGACACTGCGCCCTGATGGCAGCTTCGCTTGAGCTGGAGGTCAGCAT 2820  
QY 3052 GGCAGCTGGAGAAACAGGCCCTTAA 3075  
Db 2821 GGCAGCTGGAGAAACAGGCCCTTAA 2844

RESULT 12  
CQ845924  
LOCUS  
DEFINITION  
Sequence 4571 from Patent EP1440981.  
CQ845924  
ACCESSION  
CQ845924.1 GI:50895523  
VERSION  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,  
Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.  
Full-length human cdna  
Patent: EP 1440981-A 4571 28-JUL-2004;  
Research Association for Biotechnology (JP)  
LOCATION/Qualifiers  
1..3152  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 58.2%; Score 2673.4; DB 6; Length 3152;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 36; Indels 10; Gaps 7;  
QY 1764 TCTGATCGGGACGTCMAAGCTGACAACTGCTCTCTGCCAGCATGGAGCCAGCCAGC 1823  
Db 340 TCCCGGGGGCGCTGGCAGCTGACACGCTCTCTGCCAGCATGGAGCCAGCCAGC 399  
QY 1824 CTTCTGTCATTTGGCCATGTGTGTCTTCAACTGATGCTCTGGGAAAGTCTTTGT 1883  
Db 400 CTTCTGTCATTTGGCCATGTGTGTCTTCAACTGATGCTCTGGGAAAGTCTTTGT 459  
QY 1884 CACAGGGACTACATCTCTGGCAGACAGACCAATGCTCGGAGTGTCTGGGAG 1943  
Db 460 CACAGGGACTACATCTCTGGCAGACAGACCAATGCTCGGAGTGTCTGGGAG 519  
QY 1944 GAGCTGCACGCCCAAGGTGGATGTCTGGAGCAGCTGTGTATGCTGCACATCTCAA 2003  
Db 520 GAGCTGCACGCCCAAGGTGGATGTCTGGAGCAGCTGTGTATGCTGCACATCTCAA 579  
QY 2004 CGCTGTCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCTGCTCAAGATTGCCAGGA 2063  
Db 580 CGCTGTCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCTGCTCAAGATTGCCAGGA 639  
QY 2064 GCCTCCGCTGTGAGGGAGATCCACCTCTCTGGCCCTCTCAGAGCCAGGCCATCCA 2123  
Db 640 GCCTCCGCTGTGAGGGAGATCCACCTCTCTGGCCCTCTCAGAGCCAGGCCATCCA 699  
QY 2124 AGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTCAGCGGAGCTGGAGGGAAGT 2183  
Db 700 AGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTCAGCGGAGCTGGAGGGAAGT 759  
QY 2184 GAAACGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGA 2243  
Db 760 GAAACGGGCACTACAGCAAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGA 819  
QY 2244 ACCAAGACATCCACCGCCAAATCAGCCAAATACACAGACCCCTCCATGCCAGCCGAG 2303  
Db 820 ACCAAGACATCCACCGCCAAATCAGCCAAATACACAGACCCCTCCATGCCAGCCGAG 879  
QY 2304 AGAGCTTTCGCAAGGGCCCCAGGGCCCCCGCGCAGCTGAGGAGACAAACAGGAGGCC 2363  
Db 880 AGAGCTTTCGCAAGGGCCCCAGGGCCCCCGCGCAGCTGAGGAGACAAACAGGAGGCC 939  
QY 2364 TAAGCTCCAGCTCTCTTCCCAACAGAGCCCCCAAGCAAAAGTCTCTCTCCCTTGAC 2423  
Db 940 TAAGCTCCAGCTCTCTTCCCAACAGAGCCCCCAAGCAAAAGTCTCTCTCCCTTGAC 999

QY 2424 TTTGAGCAAGGAGTCTGGGATGTGGGAACCTTACCTCTGTCCTCCCTGGAGCCAGC 2483  
DB 1000 TTTGAGCAAGGAGTCTGGGATGTGGGAACCTTACCTCTGTCCTCCCTGGAGCCAGC 1059  
QY 2484 CCCTGCCAGAAACCCAGCTCACAGAGCGGAAGAACCGTCCGGAGCAGGAACCTGCA 2543  
DB 1060 CCCTGCCAGAAACCCAGCTCACAGAGCGGAAGAACCGTCCGGAGCAGGAACCTGCA 1119  
QY 2544 GCAGCTGGAATAGAAATATTCTCAACAGAGCTGTCCAGCAATTTCTCTGGAGGAGCA 2603  
DB 1120 GCAGCTGGAATAGAAATATTCTCAACAGAGCTGTCCAGCAATTTCTCTGGAGGAGCA 1179  
QY 2604 GGAGCAATTTCTCTGTCGCTCAGCATCGACAGCCTCTCCCTGTGCGATGACAGTGA 2663  
DB 1180 GGAGCAATTTCTCTGTCGCTCAGCATCGACAGCCTCTCCCTGTGCGATGACAGTGA 1239  
QY 2664 GAAACCATCAAGGCTCTCAAGCTCCGGGACACCTGAGCTCAGGCGTACACTCTG 2723  
DB 1240 GAAACCATCAAGGCTCTCAAGCTCCGGGACACCTGAGCTCAGGCGTACACTCTG 1299  
QY 2724 GAGCAGCAGGCGCTCAAGCTCCAGCTCGGAGACACCTGAGCTCAGGCGTACACTCTG 2783  
DB 1300 GAGCAGCAGGCGCTCAAGCTCCAGCTCGGAGACACCTGAGCTCAGGCGTACACTCTG 1359  
QY 2784 CACCGACACCCCAAGCTATTTCATATGCTGTGAAAGTCCAAATACAGTCTCTTAATGCTGA 2843  
DB 1360 CACCGACACCCCAAGCTATTTCATATGCTGTGAAAGTCCAAATACAGTCTCTTAATGCTGA 1419  
QY 2844 ACACCTGCACATCCGGGATTTCCACCGGTTCAAGTGGGAGACATGCGCACTGCGATCAG 2903  
DB 1420 ACACCTGCACATCCGGGATTTCCACCGGTTCAAGTGGGAGACATGCGCACTGCGATCAG 1479  
QY 2904 CAGCAGATCCAGCTGAGCTTTCAGCTTGGTCAACCAAGACGGGACGCTGTGCTGA 2963  
DB 1480 CAGCAGATCCAGCTGAGCTTTCAGCTTGGTCAACCAAGACGGGACGCTGTGCTGA 1539  
QY 2964 CGACATGAGGTGCCAGACTCGGGCATCGACTGCGAGTGCACACTGGGCGCTGATGGCAG 3023  
DB 1540 CGACATGAGGTGCCAGACTCGGGCATCGACTGCGAGTGCACACTGGGCGCTGATGGCAG 1599  
QY 3024 CTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGTGGAGAACAGGCGCTTAACCTGCCC 3083  
DB 1600 CTTCCCTGGAGCTGGAGGTCAAGCATGGCCAGTGGAGAACAGGCGCTTAACCTGCCC 1659  
QY 3084 TCCACCGCGCTCCACACTCGCGGAAGCAGCTTCTGCTCGGTGCGACGATGCTGCCC 3143  
DB 1660 TCCACCGCGCTCCACACTCGCGG - AAGCAGCTTCTGCTCGGTGCGACGATGCTGCCC 1718  
QY 3144 TGAACACACAGCTCAGCGCTTCCAGGGGATTGCGAGCGCTTCCAGCGCTCAGTGGGAAC 3203  
DB 1719 TG - AACAACAGCTCAGCGCTTCCAGGGATCTGCGAG - CCCCAGCTCAGCAGTGGAC 1776  
QY 3204 CAGGCGCTCGCAGCAGCAAGTGGGGCAAGCAGAAATGCTCCAGGATTTTACACCTGA 3263  
DB 1777 CAGGCGCTCGCAGCAGCAAGTGGGGCAAGCAGAAATGCTCCAGGATTTTACACCTGA 1836  
QY 3264 GCCCTGCCCAACCTGCTGAAACAA - TCCGCCACGTGAAGACAGAGAGGAGATGGC 3322  
DB 1837 GCCCTGCCCAACCTGCTGAGAAACAACTCCGCCACAGTGAAGAGACAG - AGGAGGATGGC 1895  
QY 3323 AGGAGTTACCTGGGGAACAAACAGGGATCTTTTCTGCCCTGCTCCAGTCCAGTTGG 3382  
DB 1896 AGGAGTTACCTGGGGAACAAACAGG - - - - ATCTTCTGCTCCAGTCCAGTTGG 1951  
QY 3383 CTTGACCGCTTGGATCAGTACCAATTTGTTGGCAGACAGGGGAGCAGCTTCCAGCCT 3442  
DB 1952 CTTGACCGCTTGGATCAGTACCAATTTGTTGGCAGACAGGGGAGCAGCTTCCAGCCT 2011  
QY 3443 GGGTCAGAAAGGGTGGGCGAGCCCTTCCGCCCTCACCTCCAGGCTGCTGTGAGAGTGT 3502  
DB 2012 GGGTCAGAAAGGGTGGGCGAGCCCTCGGCCCTCACCTCCAGGCTGCTGTGAGAGTGT 2071

QY 3503 CAAGTGTGTAAAGGCCCAAACTCAGGTTTCAAGTTCAGAGAACAGAGTTCAGAGGATATGCCCGC 3562  
DB 2072 CAAGTGTGTAAAGGCCCAAACTCAGGTTTCAAGTTCAGAGAACAGAGTTCAGAGGATATGCCCGC 2131  
QY 3563 CCCTAGGTTAAAGGGGCCCTCTAAACCCCTTGCCTGGCCTCACCTGGCCAGCTCACCCCT 3622  
DB 2132 CCCTAGGTTAAAGGGGCCCTCTAAACCCCTTGCCTGGCCTCACCTGGCCAGCTCACCCCT 2191  
QY 3623 TTTGGGTGTAAAGGGAAGAAATGCTCAGCCCTCGGAAAGGCTCCCTGGTGTAGAAATACACCAC 3682  
DB 2192 TTTGGGTGTAAAGGGAAGAAATGCTCAGCCCTCGGAAAGGCTCCCTGGTGTAGAAATACACCAC 2251  
QY 3683 ACTTTTCAGGTTGTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTTACGCAAGACCAAA 3742  
DB 2252 ACTTTTCAGGTTGTGTGCAACACAGGCTCTGAGTTGACCTCTGGTTTACGCAAGACCAAA 2311  
QY 3743 GAAGTGTGTAAAGTGAAGTGTCTCAGTCCCCCAGACATGTGCCCTTTGCTGTGGCTA 3802  
DB 2312 GAAGTGTGTAAAGTGAAGTGTCTCAGTCCCCCAGACATGTGCCCTTTGCTGTGGCTA 2371  
QY 3803 CCACTCTTTCCAGAGCAGAGCCCGAGCCCTTCAGGCCACAGCACTGCCCCAGACTC 3862  
DB 2372 CCACTCTTTCCAGAGCAGAGCCCGAGCCCTTCAGGCCACAGCACTGCCCCAGACTC 2431  
QY 3863 GCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTTGATGAGGATATGCTGTGACAGG 3922  
DB 2432 GCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTTGATGAGGATATGCTGTGACAGG 2491  
QY 3923 AACAGTCTGGATGGATGATCAGTCTTAAGAAAGCAGCAGAGAGAGAG - TCCGGC 3981  
DB 2492 AACAGTCTGGATGGATGATCAGTCTTAAGAAAGCAGCAGAGAGAGAGAGCTCCGGC 2551  
QY 3982 GCCCAGCCCACTATCAGTGTCCAGGTGCTGCTGCCAGAGAGGTTACTGCCAGCGGCACTTTCACATC 4041  
DB 2552 GCCCAGCCCACTATCAGTGTCCAGGTGCTGCTGCCAGAGAGGTTACTGCCAGCGGCACTTTCACATC 2611  
QY 4042 ACTGACACTCACCTGCTGCCCTGCCAGAGAGGTTACTGCCAGCGGCACTTTCACATC 4101  
DB 2612 ACTGACACTCACCTGCTGCCCTGCCAGAGAGGTTACTGCCAGCGGCACTTTCACATC 2671  
QY 4102 TGATGACTCAAGACACTTTTCATGCTGCCCTTCCAGAGGCGGCGAGGCGAGTGCAC 4161  
DB 2672 TGATGACTCAAGACACTTTTCATGCTGCCCTTCCAGAGGCGGCGAGGCGAGTGCAC 2731  
QY 4162 TGTAAGGAGTAGCAAGCAGGAGATGGGTGAAGGACACAGTCTTTCAGGCTGTCCACAT 4221  
DB 2732 TGTAAGGAGTAGCAAGCAGGAGATGGGTGAAGGACACAGTCTTTCAGGCTGTCCACAT 2791  
QY 4222 GCATGTGACTCCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCTTCCCATTTG 4281  
DB 2792 GCATGTGACTCCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCTTCCCATTTG 2851  
QY 4282 CCCAGCTTACCTTCTCCAGGGAGTACTCAGGACTCAGTAGCATTAATCAGC 4341  
DB 2852 CCCAGCTTACCTTCTCCAGGGAGTACTCAGGACTCAGTAGCATTAATCAGC 2911  
QY 4342 TGTCAATCGTCAGGGGTGTCTAGCTCAACCTCTCTGGGCGAGGCGCGGAGACT 4401  
DB 2912 TGTCAATCGTCAGGGGTGTCTAGCTCAACCTCTCTGGGCGAGGCGCGGAGACT 2971  
QY 4402 CCGTGGGAGAGCTCATTTCCCACTTTGCGCAAGACAGCTTTTGTCCAGCTGTCCACAT 4461  
DB 2972 CCGTGGGAGAGCTCATTTCCCACTTTGCGCAAGACAGCTTTTGTCCAGCTGTCCACAT 3031  
QY 4462 GAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCCGAGCAGCATTAAGAACTGCGAGCTT 4521  
DB 3032 GAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCCGAGCAGCATTAAGAACTGCGAGCTT 3091  
QY 4522 GGTACTCGAGTCTGGTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTGA 4581  
DB 3092 GGTACTCGAGTCTGGTGTGAGAGACTCTTTGTAGCAATAAAGTTTGGGGTGTGA 3151  
QY 4582 C 4582



QY 2904 CAGCAGATCCAGCTGAGCCTTACAGTTGGTACCAAGACGGGAGCCTTTCGCTA 2963  
Db 1480 CAGCAGATCCAGCTGAGCCTTACAGTTGGTACCAAGACGGGAGCCTTTCGCTA 1539  
QY 2964 CGACATGGAGGTGCGAGACTCGGCGATCGAGCTGAGTGACACTGGCCCTGATGCGAG 3023  
Db 1540 CGACATGGAGGTGCGAGACTCGGCGATCGAGCTGAGTGACACTGGCCCTGATGCGAG 1599  
QY 3024 CTTCCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGAGAACAGGGCCCTAACCCCTGCC 3083  
Db 1600 CTTCCGCTGGAGCTGGAGGGTCAAGCATGGCCAGCTGAGAACAGGGCCCTAACCCCTGCC 1659  
QY 3084 TCCACCGCCGCTCCACACTGCGCGAAGAGAGCCTTCTGCTCGGTGCGACGATGCTGCC 3143  
Db 1660 TCCACCGCCGCTCCACACTGCGCGG-AAAGAGCCTTCTGCTCGGTGCGACGATGCTGCC 1718  
QY 3144 TGAAGACACAGGCTCAGCGCTTCCAGGGGATTCGACGCCCCCGGCTCAGAGTGGGAC 3203  
Db 1719 TG-AAACACAGGCTCAGCGCTTCCAGGGATTCGCGAG-CCCCCGCTCAGAGTGGGAC 1776  
QY 3204 CAGGGCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCTCCAGAGATTTTCAACCTGA 3263  
Db 1777 CAGGGCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCTCCAGAGATTTTCAACCTGA 1836  
QY 3264 GCCCTGCCACCTCTGTGAAAAAACA-TCCGCCACGTGAAGACAGAGAGGATGGC 3322  
Db 1837 GCCCTGCCACCTCTGTGAAAAAACAATCCGCCACGTGAAGACAG-AGGAGGATGGC 1895  
QY 3323 AGGAGTTACCTGGGGAAACAAACAGGATCTTTTCTGCCCCCTCCAGTCAGTGG 3382  
Db 1896 AGGAGTTACCTGGGGAAACAAACAGG---ATCTTCTGCCCCCTCCAGTCAGTGG 1951  
QY 3383 CCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGGGGAGAGCAGCTTCCAGCCT 3442  
Db 1952 CCTGACCCGCTTGGATCAGTGACCATTTGTGGCAGACAGGGGAGAGCAGCTTCCAGCCT 2011  
QY 3443 GGGTCAGAGGGGTGGGAGCCCTTGGCCCCCTCAACCTCCAGGCTGTGTGAGAGTGT 3502  
Db 2012 GGGTCAGAGGGGTGGGAGCCCTTGGCCCCCTCAACCTCCAGGCTGTGTGAGAGTGT 2071  
QY 3503 CAAGTGTGTAAGGGCCCAACACTCAGGTTCAAGTCAGTGCAGAACAGGTCAGAGGTATGCCGC 3562  
Db 2072 CAAGTGTGTAAGGGCCCAACACTCAGGTTCAAGTCAGTGCAGAACAGGTCAGAGGTATGCCGC 2131  
QY 3563 CCGTAGGTTAAGGGGGCCCTTAAACCCCTTGGCTGGCCCTCACTGGCCAGCTCACCCCT 3622  
Db 2132 CCGTAGGTTAAGGGGGCCCTTAAACCCCTTGGCTGGCCCTCACTGGCCAGCTCACCCCT 2191  
QY 3623 TTTGGGTGTAAGGGGAAAGAAATGCTGACCTGGGAAAGGCTCCCTGTGTGAGTAACACAC 3682  
Db 2192 TTTGGGTGTAAGGGGAAAGAAATGCTGACCTGGGAAAGGCTCCCTGTGTGAGTAACACAC 2251  
QY 3683 ACTTTTCAGGTTGTGCAACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACCAAA 3742  
Db 2252 ACTTTTCAGGTTGTGCAACAGCTCTGAGTTGACCTCTGTTTCAGCCAGGACCAAA 2311  
QY 3743 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCCAGACATGTGCCCTTTGCTGTGGCTA 3802  
Db 2312 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCCAGACATGTGCCCTTTGCTGTGGCTA 2371  
QY 3803 CCACCTTTCCAGACAGCAGGCCCCAGCCCCCTTACGGCCCCAGCAGCTGCCAGACTC 3862  
Db 2372 CCACCTTTCCAGACAGCAGGCCCCAGCCCCCTTACGGCCCCAGCAGCTGCCAGACTC 2431  
QY 3863 GCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGAGGATATGCTTCAGAG 3922  
Db 2432 GCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGAGGATATGCTTCAGAG 2491  
QY 3923 AACAGTCTGTGGATGGACATCATCAGTCTGAAGGAAAGCAGCAGAGAGAGACG-TCCGGC 3981  
Db 2492 AACAGTCTGTGGATGGACATCATCAGTCTGAAGGAAAGCAGCAGAGAGAGACGCTCCGC 2551  
QY 3982 GCCCCAGCCCCCATATCAGTGTCCAGGGTGTGTTTCCCCAGAGCAGCAGCTCAGCATCAC 4041

Db 2552 GCCCCAGCCCCCATATCAGTGTCCAGGGTGTGTTTCCCCAGAGCAGCTCAGCATCAC 2611  
QY 4042 ACTGACACTACCTCGCTGCCCTGCCAGAGGGTACTGCCAGACGGCACTTTTGCACCTC 4101  
Db 2612 ACTGACACTACCTCGCTGCCCTGCCAGAGGGTACTGCCAGACGGCACTTTTGCACCTC 2671  
QY 4102 TGATGACCTCAAGACACTTTTCATGGCTGCCCTCTGGCAGAGGGCAGGGCAGGCACTGACAC 4161  
Db 2672 TGATGACCTCAAGACACTTTTCATGGCTGCCCTCTGGCAGAGGGCAGGGCAGGCACTGACAC 2731  
QY 4162 TGATGAGCATAGCAAGCAGGAGATGGGGTGAAGGGACACAGTCTTTGAGCTGTCCACAT 4221  
Db 2732 TGATGAGCATAGCAAGCAGGAGATGGGGTGAAGGGACACAGTCTTTGAGCTGTCCACAT 2791  
QY 4222 GCATGTGACCTCTCAAACTCTTCCAGATTTCTTAAGAAATAGCACCCCTTCCCCCATTTG 4281  
Db 2792 GCATGTGACCTCTCAAACTCTTCCAGATTTCTTAAGAAATAGCACCCCTTCCCCCATTTG 2851  
QY 4282 CCCAGCTTACGCTTCTTCCAGAGGAGTACTCAGGACTCAGTAGCATTAATAATCAGC 4341  
Db 2852 CCCAGCTTACGCTTCTTCCAGAGGAGTACTCAGGACTCAGTAGCATTAATAATCAGC 2911  
QY 4342 TGTGAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTTGGGGCAGGGGACGCCGAGACT 4401  
Db 2912 TGTGAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCTTGGGGCAGGGGACGCCGAGACT 2971  
QY 4402 CCGTGGGAGAAAGCTCATTTCCACATCTTGGCAACAGACAGCCCTTTCAGCTGTCCACATT 4461  
Db 2972 CCGTGGGAGAAAGCTCATTTCCACATCTTGGCAACAGACAGCCCTTTCAGCTGTCCACATT 3031  
QY 4462 GAGTCAGACTGCTCCCGGAGAGAGCCCCGGCCCCCAGACACATAAAGAACTGAGGCTT 4521  
Db 3032 GAGTCAGACTGCTCCCGGAGAGAGCCCCGGCCCCCAGACACATAAAGAACTGAGGCTT 3091  
QY 4522 GGTACTGACAGCTGCGGTGTAGAGAACTCTTTGTAAAGCAATAAGTTGGGGGTGATGA 4581  
Db 3092 GGTACTGACAGCTGCGGTGTAGAGAACTCTTTGTAAAGCAATAAGTTGGGGGTGATGA 3151  
QY 4582 C 4582  
Db 3152 C 3152

RESULT 14  
AF143094  
LOCUS AF143094 2829 bp mRNA linear ROD 20-MAY-1999  
DEFINITION Mus musculus NF-kappaB inducing kinase (Nik) mRNA, complete cds.  
ACCESSION AF143094  
VERSION AF143094.1 GI:4877962  
KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE 1 (bases 1 to 2829)  
AUTHORS Shinkura,R., Kitada,K., Matsuda,F., Tashiro,K., Ikuta,K.,  
Suzuki,M., Kogishi,K., Serikawa,T. and Honjo,T.

TITLE Alymphoplasia is caused by a point mutation in the mouse gene  
encoding Nf-kappa b-inducing kinase

JOURNAL Nat. Genet. 22 (1), 74-77 (1999)

MEDLINE 99251583

PUBMED 10119865

REFERENCE 2 (bases 1 to 2829)

AUTHORS Shinkura,R., Kitada,K., Matsuda,F., Tashiro,K., Ikuta,K.,

Suzuki,M., Kogishi,K., Serikawa,T. and Honjo,T.

TITLE Direct Submission

JOURNAL Submitted (14-APR-1999) Medical Chemistry, Kyoto University,

Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8315, Japan

FEATURES Location/Qualifiers

source

1. .2829

/organism="Mus musculus"

/mol\_type="mRNA"



gene	/strain="C57BL/6"		Db			541	GATGAGTCTCCACTAGGCAACCTCTATGCCAGAAATGTCTCCAGTTCACCAAGCCTCTG		600
	/db_xref="taxon:10090"								
CDS	/chromosome="11"		Qy			832	AAGGAACAGGCGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACCGCGCGCT		891
	/map="between D11Mit331 and D11Mit360"								
	1..2829		Db			601	GGGGACAGGCGCTTGGCCACCTGTGCTTTAAGAAACAAGATGAAGGCTTGGCAGCCGGTA		660
	/gene="Nik"								
	1..2829		Qy			892	CTGCGCTCGATCAGAACTCCACAAACTGATCAGCCCTTGCATGTCTGAACACACGCTGTGG		951
	/gene="Nik"								
	/note="Nik"		Db			661	CTGCGCTCGACCAAACTCCACAACTGATCAGCCCTTGCATGTCTGAATGTCTAAACACGCTGTGG		720
	/codon_start=1								
	/product="NF-kappaB inducing kinase"		Qy			952	AAACTGCACCAACCCCAAGGACGGAGGCCCTGCGCCCTGCCACGACGCCCTTCCCTTAT		1011
	/protein_id="AAD31512.1"								
	/db_xref="GI:4877963"		Db			721	AAACTCCACCAACCCCAAGGCGCCCGCGGCCACCGACTACACCCCTTCCCTTAT		780
	/translation="WAVMEVACPGTPGSAVGQOKELAKAKEKTQSLGKQSCIFKLEA								
	VEKSPVFCGWKILNDVITKGTADSGEGPPAISIIAQAECEENSOEFTFSERIFI		Qy			1012	AGCAAGTGCCTCATCCTTCCCATTCACCCCTCTCCAGCCCTGGAACCTCACCCCTCTG		1071
	AGSQYSSQESDQIPNNVAHATEGMARVCRRKRGKARKRRKSKSLAQGVA								
	LAKPLRPEQESCTIPVDESDPLNLYARNVSQTPKPLGGFGHJLCKFKQDEGLR		Db			781	AGCGAATGCCCCATCTTCCCATTTACCCCTTGGAGCCCTGGAACCTTATATGCTG		840
	PVLPRPEHLKLSPLQCNLHWKHPQATPRPHPTFPFYSMPHPFFYPLEPWK								
	PYMLDSALDKLAGVGSGRPLPFPPLHLSQAHGDSQKPLPGPHLESSCPSRGALEKVP		Qy			1072	GAGT---CCTTCCTGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGCCTGACCCA		1128
	VEEYLVAHQSSQAHSLASLAKTSSGSAKLQRLGPETEDNEGVLITLKLKPDV								
	VEYREVHMTHPQVRGSGFGEVHRMDKOTGFQCAVKVRLVFRVEELVACAGLS		Db			841	GACTCTGCGCTCTGACAAACTAGCCGCTGTCAGCGGCCACGGCCTCTGCCTGGCCCA		900
	SPRIYPLVGAVERGPWNIEMELLESGSLGOLIKOMGCLPEDRALYLYLOALRGLEYL								
	HTRRLHGDVADNVLSSDGSRAALCDFGHALCLQPDGLGKSLTGDYIPGTETHWA		Qy			1129	---CACCTTGAGCAAACTGGCCTGTGTAGACAGTCCAAAAGCCCTGCCTGGCCCAACCTG		1185
	PEVMWGPKDAKVDIWSCCMMLHMLNGCHFWTQYFRGPLCLIKIASPEPPREIPDQA								
	APLTAQATQEGURKEPVRHASAMELRKVKALQEVGLKSPWKGEYKPRPPDQA		Db			901	CGCATCTAAGCCAACTGGCCCATGGAGACAGTCAAGAGCCGCTGCCTGGGCCCAACCTG		960
	TCHQTLPTPRENPAPKANTDGAPEPQPLPEPPEPSKAPALNLSKEESGTWEPLPL								
	SSLDPATKPGSPDPRATLPELELOQLEIELEFLNLSQPFSLERQEOILSCLSDSL		Qy			1186	GAGCCCAAGTGCCTCTCTGCTGCTGCCATGAGAAAGTCTTCTGTGAGGAATACCTAGTG		1245
	SLSDSEKNPSKASOSRDTLSSGVHSMNSOAEARTCSCTALARGRPDTIPSYFNGV								
	KVQISLNGELHLIREFHRVKVDIATGISSQIPATAPSLVTKDQGPVCYDMEVPDGG		Db			961	GAGTCCAGCTGCCCGCTCTCGGGGTGCCCTAGAAAAGGTTCGCGTGAGGAATACCTGGTG		1020
	IDLQCTLAPDGSFATWRVKHGLENNR"								
ORIGIN	Query Match		Qy	44.7%; Score 2054.6; DB 10; Length 2829;		1246	CATGCTCTGCAAGGACGCTGAGCTCAAGCCAGGCCCAAGCTGACAGCTGACAGCTGGCCAAAG		1305
	Best Local Similarity			83.5%; Pred. No. 0;					
	Matches 2378; Conservative		Db	0; Mismatches 444; Indels 27; Gaps 3;		1021	CATGGCTCCAAAGGAAGTGTGAGCTCAGGCCAGGCCCAAGCTTGGCCAGCTTGGCTTAAG		1080
	232 ATGGCAGTGAATAATGCCCTGCCAGGTGCCCTGCTCAGCAGTGGGGCAGAGAAG		Qy	291		1306	ACCTGGGACAGCAGGGGCTCCAGATCCCGGGAGCCAGCCCAAACTGAGGACAAACGAG		1365
	1 ATGGCGGTGATGAAATGGCTTGCCTGGGCACTCTCTGGGTACAGCTCGGGCAGAGAAG		Db	60		1081	ACATGGTCTCTCGGGAAGCGCCAAAGCTGCAGAGGCTCGGCCCGGAAACTGAGGACAAACGAG		1140
	292 GAATCCCAAGCAAGAGAGAGCGCGCCACTTGGGGAAGAAACAGAGCTCCGTCTAC		Qy	351		1366	GGTGTCTGTCTCACTGAGAACTCAAGCCAGTGGATATGAGTACCGAGAGAGTCCAC		1425
	61 GAGCTTGCCAAAGGAGAGACACAGTCACTGGGGAAGAACAGAGCTGCATCTTC		Db	120		1141	GGGTCTCTGCTTACTGAGAAACTCAAGCCAGTGGATATGAGTATCGAGAAGAGGTCCAC		1200
	352 AAGCTTGAGCGCTGGAGAAGCGCTGTGTTCTCGGAAAGTGGGAGATCTCGAATGAC		Qy	411		1426	TGGGCCACGACACGAGCTCCGCTGGGACAGAGGCTCTTCGAGAGAGTGCACAGATGGAG		1485
	121 AAGCTTGAGCGCTGGAGAAGCGCTGTGTTCTCGGAAAGTGGGAGATCTCGAATGAC		Db	180		1201	TGGATGACACACAGCTCGGGTGGGACAGAGGCTCTTCGCGAGAGTCCACAGAAATGAAG		1260
	412 GTGATTAACAAGGACACAGCAAGAGAGCTCCGAGGACAGGCCAGCTGCCATCTCTATC		Qy	471		1486	GACAAAGACAGTGGCTTCCAGTGGCTGTCAAAAAGGTGCGCTGGAAGTATTTCCGGCA		1545
	181 GTGATCAACAAAGGACACAGCCAGAGCGCTCTGAGGAGGACACCCGGCCATCTCCATC		Db	240		1261	GACAAAGACAGAGCTTCCAGTGGCTGTCAAAAAGGTGACGACTCGAGGTGTTTCCGGGTA		1320
	472 ATGCGCCAGGCTGAGTGTGAGATAGCCAAAGTTCAGCCCAACCTTTTCAGAAAGCATT		Qy	531		1546	GAGGAGCTGATGGCATGTGCAGGATTCACCTCACCCAGAAATGTCCTTCTGTATGAGCT		1605
	241 ATGCGCCAGGCTGAAATGTGAGAAATAGCCAAAGATTCAGCCCAACCTTTCTCAGAGCGCAT		Db	300		1321	GAGGAACTAGTGGCGCTGTGCTGGGTCTGAGCTCGCCCAAGAAATCGTCTCTATGAGCT		1380
	532 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT		Qy	591		1606	GTGAGAGAGGCGCTTGGGTCAACATCTTCATGAGCTGCTGGAAGTGGCTCCCTGGGC		1665
	301 TTTCATCGCGGGTGCACAGCTACAGCCAGTCTGAGAGTCTCGATCAATATCCCAACAAT		Db	360		1381	GTGAGAGAGGCGCTGGGTGAACATCTTCATGGAATCTGTAGAAAGTGGCTCGCTGGGT		1440
	592 GTGGCCCATGTACAGAGGCAAAATGGCCCGTGTGTGTGGAAGGAAAGCGTCCGACG		Qy	651		1666	CAGCTGTCAAGGAGCAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACTTGGGCCAG		1725
	361 GTGGCCCATGTCAACTGAAGGCAAAATGGCCCGTGTGTGTGGAAGGAAAGCGTCCGCGC		Db	420		1441	CAGCTCATAAAGCAAAATGGGCTGTCTGCCAGAAAGCCGAGCCCTTTACTACTTGGGCCAG		1500
	652 AAAGCCCGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGAGGAGTGGCC		Qy	711		1726	GCCCTGGAGGCTTGGAAATACCTCACTCAAGAGGATTCCTCATGGAAGTCAAGCT		1785
	421 AAAGCCCGAAGAAACGTAAGGAAGAGGTGCGAAGTCACTGGCCAGGAGGAGTGGCC		Db	480		1501	GCCCTGGAGGGGCTGGAGTACCTTCCACACAGCCAGGAATTCGTATGGCGATGTCAAAGCT		1560
	712 TTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAGCTGACCATCCCTGAGTGGAGG		Qy	771		1786	GACAACTGCTCTCTGTCCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGGCCATGTCT		1845
	481 TTAGCCAAAGCCCTGCCAGAACCCCTGAGCAAGAGAGCTGTACCATCCCAAGTACAGGA		Db	540		1561	GACAACTGCTCTCTGTCCAGTGTGGAAGCCGAGCGGCCCTCTTGCAGCTTTGGCCACGCC		1620
	772 GATGAGTCTTCAACTCGGCGCCGCCATATGTTGTAGAAACACCCCGCAGGTTTCAACAAAGCCTCTG		Qy	831		1846	GTGTGTCTTCAACTGATGGCTGGGAAAGTCTTGTGCTCACAGGGGACTACATCCCTTGGC		1905



Db 1621 TTGTGCTGCAACCTGACGGCTAGGGAAATCTCTTGCTCA CAGGGGACTACATCTCTGGC 1680  
Qy 1906 ACAGAGACCCACATGGCTCCGAGGTGTGTGGGACGAGCTGCGAGCGCAAGTGTGAT 1965  
Db 1681 ACGGAGACCCACATGGCACCAAGAGTGTGTGATGGAAAGCCCTGGATGCCAAGGTGGAC 1740  
Qy 1966 GTCTGGAGCAGCTGTGTATGATGTGCATGTCTAAACGGGTGCGACCCCTGTGACTCAG 2025  
Db 1741 ATCTGGAGCAGCTGTGTATGATGTCTCAATGTCTCAACGGGTGCGACCCCTGTGACTCAG 1800  
Qy 2026 TTCTTTCCGAGGCGCTCTGCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATC 2085  
Db 1801 TACTTCCAGGCGCGCTTTGTCTCAAGATTGCCAGGAGCTTCAACCGATCAGGGAGATC 1860  
Qy 2086 CCACCTCTCTGCGCCCTCTCACAGCCAGGCGCATCAAGAGGGGCTGAGAAAGAGAGCC 2145  
Db 1861 CCACCTTCTGGCACCCCTCACAGCCAGGCGCATCAAGAGGGGCTGAGAAAGAGAGCC 1920  
Qy 2146 ATCCACCGCTGTCTGACGCGAGCTGGGAGGGAGGTGAACCGGGCACTACAGCAAGTG 2205  
Db 1921 GTCCACCGAGCATCTGCCATGAGCTTCGGAGGAAAGTGGCAAGGCACTACAGCAAGTG 1980  
Qy 2206 GGAGGTCTGAGAGCCCTTGGAGGGGAGATATAAGAACCAAGACATCCACCCGCAAT 2265  
Db 1981 GGAGGTCTGAAAGCCCTTGGAAAGGGAGATATAAGAACCAAGACCTTCCACCCCAAGAC 2040  
Qy 2266 CAAGCCAAATTACCAACAGACCTCTCATGCCAGCGAGAGAGCTTTCGCCAAGGGCCCCA 2325  
Db 2041 CAAGCCACCTGCCACAGACCTTACTTACTCGCGGAGAGA----- 2081  
Qy 2326 GGGCCCGCGGCTGAGGAGACAAAGCTCTCTCTCTGACTTTGAGCAAGGAGGCTCTGGG 2445  
Db 2082 --GAACCCACAGCAAGGCCAACACAGACGGGGCTCTGAGCCTCAGCCTCTCTACCG 2139  
Qy 2386 CCAGAGCCCCCAGAGGCCAAAGTCTCTCTCTCTGACTTTGAGCAAGGAGGCTCTGGG 2445  
Db 2140 CCAGAACCCAGAACCGAGCAAAAGCGCCAGCCCTGAACTTGAGCAAGGAGGAGTCTGGC 2199  
Qy 2446 ATGTGGGAACCTTACTCTCTCTCTCTCTGAGCGAGCGCCCTGCGAGAAACCCAGCTCA 2505  
Db 2200 ACATGGGNAACCTTGCTCTCTCTCTGAGCCAGCCATGCGCAAGGCCCTCAGCTTC 2259  
Qy 2506 CCAGAGGGGAAGCAACCGTCCCGAGCAGAACTGACGAGCTGGAATAAGAAATTATTC 2565  
Db 2260 CCAGACCGAGGGCAACCTTGCCAGAGCTGAGCTACAGCAACTGGAGATAGAACTGTTT 2319  
Qy 2566 CTCAACAGCTGTCCAGCCATTTTCTCTGAGAGCAGGAGCAAAATTTCTCTCTGCGCTC 2625  
Db 2320 CTCAACAGCTGTCCAGCGCTTCTCTCTGAGGAAACAGGAAACAAATCTCTCTCTGCGCTC 2379  
Qy 2626 AGCATCGACAGCTCTCTCTCTGATGACAGTGAAGAAACCCATCAAGGCGCTCTCAA 2685  
Db 2380 AGCATCGACAGCTCTCTCTCTGATGACAGTGAAGAAATCCATCGAAGGCTCTCAG 2439  
Qy 2686 AGCTCGGGGACACCTTGAGCTCAGGGGTACACTCTCTGGAGCAGCGCGGAGGCTCGA 2745  
Db 2440 AGCTCAGGGACACCTTGAGTCTGGGTGCACTCTTTGGAAACAGCAAGCTGAGGCAAGA 2499  
Qy 2746 AGCTCAGCTGGAACATGTTGCTGGCGGGGGCGGCCACCGACACCCCAAGCTATTTC 2805  
Db 2500 ACCTGCACTGTCAGCAGCGCTGGCGGGGGGGGGCTACTGACATCCCGAGCTACTTC 2559  
Qy 2806 AATGGTGAAGTCCAAATACAGTCTCTTAATGTGAACACCTGACATCCGGGAGTTC 2865  
Db 2560 AACGGGTGCAAGGTCCAGATCCAGTCTCTCAATGGCGAAACCTTGCATATCCGGGAATTC 2619  
Qy 2866 CACCGGTGCAAGTGGGAGACATGCCACTGGCATCAGCAGCCAGATCCCAAGCTGCGAGCC 2925  
Db 2620 CACCGGTGCAAGTGGGAGACATTTGCCACCGCATCAGCAGCCAGATCCCAAGCTGCGAGCT 2679  
Qy 2926 TTACGCTTGTGTCACAAAGACGGGAGCTCTTCTGCTTACGACATGGAGGTGCCAGACTCG 2985  
Db 2680 TTACGCTTGTGTCACAAAGATGGACAGCTCTTCTGCTTATGACATGGAGGTGCCAGACTCG 2739

Qy 2986 GGCAATCGACCTGCACTGATGCGCCCTGATGCGAGCTTTCGCTTGAGCTGGAGGCTC 3045  
Db 2740 GGCAATCGACCTGCACTGATGCGCCCTGATGCGAGCTTTCGCTTGAGCTGGAGGCTC 2799  
Qy 3046 AAGCATGGCAGCTGGAGAACAGCCCTA 3074  
Db 2800 AAGCATGGTCACTGGAGAACCCGACCTA 2828  
RESULT 15  
A66647  
LOCUS A66647 2631 bp DNA linear PAT 29-MAR-1999  
DEFINITION Sequence 3 from Patent WO9737016.  
ACCESSION A66647  
VERSION A66647.1 GI:4538139  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 2631)  
AUTHORS Wallach,D., Malinin,N., Boldin,M., Kovalenko,A. and Mett,I.  
TITLE MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR  
PREPARATION AND USE  
JOURNAL Patent: WO 9737016-A 3 09-OCT-1997;  
YEDA RES & DEV (IL)  
COMMENT Other publication AU 2175597 19971022.  
FEATURES  
Location/Qualifiers  
source 1..2631  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

Query Match 40.3%; Score 1851.2; DB 6; Length 2631;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;  
Qy 2099 CCCCTCTCACAGCCAGCGCCATCCAGAGGGGCTGAGAAAGAGCCATCCACCGCGTGT 2158  
Db 1 CCCCTCTCACAGCCAGCGCCATCCAGAGGGGCTGAGAAAGAGCCATCCACCGCGTGT 60  
Qy 2159 CTGACGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 2218  
Db 61 CTGACGCGGAGCTGGGAGGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGA 120  
Qy 2219 GCGCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 2278  
Db 121 GCGCTTGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACC 180  
Qy 2279 ACCAGACCTCCATGCCCGAGAGAGCTTTTCGCAAGGGCCCGGGCCCGGCCAG 2338  
Db 181 ACCAGACCTCCATGCCCGAGAGAGCTTTTCGCAAGGGCCCGGGCCCGGCCAG 240  
Qy 2339 CTGAGGAGACAAACAGCGAGAGCCCTTAAGCTTCCAGCTCTCTCTCCACAGAGCCCGCCAG 2398  
Db 241 CTGAGGAGACAAACAGCGAGAGCCCTTAAGCTTCCAGCTCTCTCTCCACAGAGCCCGCCAG 300  
Qy 2399 AGCCAAACAAAGTCTCTCTCTTGAATTGAGCAAGGAGGAGTCTGGGATGAGGAAACCT 2458  
Db 301 AGCCAAACAAAGTCTCTCTCTTGAATTGAGCAAGGAGGAGTCTGGGATGAGGAAACCT 360  
Qy 2459 TACCTCTGCTCTCTCTGGAGCGAGCCCTTCCAGAAACCCAGCTCCACAGAGCGGAAAG 2518  
Db 361 TACCTCTGCTCTCTCTGGAGCGAGCCCTTCCAGAAACCCAGCTCCACAGAGCGGAAAG 420  
Qy 2519 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCCTCAACAGCTGT 2578  
Db 421 CAACCGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTATTCCTCAACAGCTGT 480  
Qy 2579 CCAGGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 2638  
Db 481 CCCAGCCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCC 540

QY	2639	TCTCCCTCTCGATGACAGTGAAGAAACCCATCAAAAGCCCTCTCAAAAGTCTCGCGGACA	2698
Db	541	TCTCCCTCTCGATGACAGTGAAGAAACCCATCAAAAGCCCTCTCAAAAGTCTCGCGGACA	600
QY	2699	CCCTGAGCTCAGCGGTACACTCTGGAGCAGCAGCGCGAGCTCGAGCTCCAGCTGGA	2758
Db	601	CCCTGAGCTCAGCGGTACACTCTGGAGCAGCAGCGCGAGCTCGAGCTCCAGCTGGA	660
QY	2759	ACATGCTGCTGCGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	2818
Db	661	ACATGCTGCTGCGCCCGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAG	720
QY	2819	TCCAAATACAGTCTCTTAATGGTGAACACTCTGACATCCGGAGTTCACCGGGTCAAG	2878
Db	721	TCCAAATACAGTCTCTTAATGGTGAACACTCTGACATCCGGAGTTCACCGGGTCAAG	780
QY	2879	TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTGGTCA	2938
Db	781	TGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTGGTCA	840
QY	2939	CCAAAGACGGGAGCGCTGTTCTGCTACGACATGAGGTGCGACATCGGGCATTCGACCTGC	2998
Db	841	CCAAAGACGGGAGCGCTGTTCTGCTACGACATGAGGTGCGACATCGGGCATTCGACCTGC	900
QY	2999	AGTGACATCGGCGCTGATGGAGCTTCGCTTGGAGCTGGAGGTCAAGCATGGCCAGC	3058
Db	901	AGTGACATCGGCGCTGATGGAGCTTCGCTTGGAGCTGGAGGTCAAGCATGGCCAGC	960
QY	3059	TGGAGAACAGGCGCTAAACCTTCCCTCCACCGCGCTCCACACTGCCGGGAAAGCAGCT	3118
Db	961	TGGAGAACAGGCGCTAAACCTTCCCTCCACCGCGCTCCACACTGCCGGGAAAGCAGCT	1020
QY	3119	TCTGCTCGGTGACGATGCTGCTTGAAGAAACAGGCTTCAGCTTCAGGGGATTT	3176
Db	1021	TCTGCTCGGTGACGATGCTGCTTGAAGAAACAGGCTTCAGCTTCAGGGGATTT	1080
QY	3177	GCAGCGCCCGGCTCA--CAGTGGGAAACAGGGCTCG----CAGCAAGAGTGGG	3229
Db	1081	NCCAGCGCCCGGCTCARCAGNTGGGAAACAGGGCTTCGNCAGCAGGAGTGGG	1140
QY	3230	GCAAGCAGATGCTCCAGGATTTACACCTTGAGCCCTGCCCCAC-----CCTGCTGA	3283
Db	1141	CAGCAGAGATGCTCCAGGATTTACACCTTGAGCCCTGCCCCAC-----CCTGCTGA	1200
QY	3284	AAAAACATCGCAGTGAAGACAGAGGAGATGGC-----AGAGTTACCTGGGAA	3339
Db	1201	AAAAATNCCCGCACGTGAAGACAGAGGAGATGGC-----AGAGTTACCTGGGAA	1260
QY	3340	ACAAACAGGATCTTTTCTGCTGCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG	3395
Db	1261	ACAAACAGGATCTTTTCTGCTGCTGCTCCAGT--CGAGTTGGCTGA--CCCGCTTG	1320
QY	3396	GATCAGTGACATTTGTTGGCAGA-CAGGGGAGAGAGCTTCAGCTGGGTGAGAAGG	3454
Db	1321	ANTCAGTGACATTTGTTGGCAGANACAGGGAGAGAGCTTCAGCTGGGTGAGAAGG	1380
QY	3455	GTGGGAGAGCTTCGGCGCTTACCT--CCAGCTGCTG--AGAGTTCAGTGTGA	3512
Db	1381	GTGGGAGAGCTTCGGCGCTTACCT--CCAGCTGCTG--AGAGTTCAGTGTGA	1440
QY	3513	AGGCGCCCAACTC--AGGTTGAGTGAGAACAGGT--CAGCAGGATGTCGCCCGCTAG	3568
Db	1441	AGGCGCCCAACTC--AGGTTGAGTGAGAACAGGT--CAGCAGGATGTCGCCCGCTAG	1500
QY	3569	GTTAA----GGGGCGCTCTAAACCCCTTGGCTGCTCA--CCTGGCAGCTCA--CCC	3620
Db	1501	GTTAAANNNGGGCGCTCTAAACCCCTTGGCTGCTCA--CCTGGCAGCTCA--CCC	1560
QY	3621	CTTTTGGGTGAGGGGAAAGAACTGCTGACCTGGGAAGCT--CCCTGGTGAATACAC	3679
Db	1561	CTTTTGGGTGAGGGGAAAGAACTGCTGACCTGGGAAGCT--CCCTGGTGAATACAC	1620

Search completed: June 10, 2005, 14:23:09  
Job time : 13081.6 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 22:58:14 ; Search time 1562.09 Seconds  
(without alignments)  
17417.120 Million cell updates/sec

Title: US-09-155-676B-6

Perfect score: 4596

Sequence: 1 agcgggggactgtgcgctg.....gatgacaaatgttaaaaaa 4596

Scoring table:  
IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4596	100.0	4596	2	AAV03326 DNA seque
2	4578.4	99.6	4596	5	Aas79490 DNA encod
3	4578.4	99.6	4596	6	ABK51171 CDNA enco
4	4578.4	99.6	4596	10	ADe85075 Farnesyl
5	4578.4	99.6	4596	13	ADR25114 Breast ca
6	4440	96.6	4684	11	ACN44439 Human mRN
7	3149.6	68.5	3156	2	AAV73917 Human NIK
8	3149.6	68.5	3156	2	AAV71603 Human NF-
9	3149.6	68.5	3156	2	AAV69285 Human NF-
10	2839.2	61.8	2844	2	AAx87842 NF-kB ind
11	2673.4	58.2	3152	12	ADQ67410 Novel hum
12	2042.6	44.4	2828	11	ACN44437 Mouse mRN
13	1851.2	40.3	2631	2	AAV03324 Clone 10
14	1556.2	33.9	73967	11	ACN44438 Human gen
15	1474.4	32.1	2760	5	AAS79491 DNA encod
16	1304	28.4	1385	4	AAH99262 Human pro
17	1241	27.0	2070	12	ADP28582 Human sec
18	627.6	13.7	722	5	AAS79492 DNA encod
19	480.2	10.4	42566	11	ACN44436 Mouse gen
20	476	10.4	476	10	ADF79802 Leukaemia

C	21	349.4	7.6	2475	11	ADM03069	Adm03069 Human cDN
	22	298.2	6.5	362	3	AAc02585	Aac02585 Human sec
	23	291.6	6.3	6673	6	ABQ67111	Abq67111 Human ang
	24	267.4	5.8	499	9	ACH15373	Ach15373 Human adu
C	25	243.6	5.3	6673	6	ABQ67112	Abq67112 Human ang
	26	171.8	3.7	500	12	ACH75442	Ach75442 Human gen
	27	166.8	3.6	293	12	ACH89170	Ach89170 Human gen
	28	89.8	2.0	707	12	ADH56283	Adh56283 Human MAP
	29	88.8	1.9	888	12	ADH56286	Adh56286 Human MAP
	30	88.8	1.9	900	12	ADH56282	Adh56282 Human MAP
	31	88.8	1.9	973	12	ADH56281	Adh56281 Human MAP
	32	88.8	1.9	1971	12	ADH56272	Adh56272 Human MAP
	33	88.8	1.9	2032	12	ADN61533	Adn61533 Human KPP
	34	88.8	1.9	2154	12	ADH56270	Adh56270 Human MAP
	35	88.8	1.9	2348	6	AAc34715	Aac34715 Human MEK
	36	88.8	1.9	2348	12	ADN71931	Adn71931 Human kin
	37	88.8	1.9	2488	13	ADR39801	Adr39801 Human kin
	38	88.8	1.9	3150	12	AD126096	Adi26096 Human cDN
	39	88.8	1.9	3286	3	AAc76868	Aac76868 Human ORF
	40	88.8	1.9	3295	13	ACN40256	Actn40256 Tumour-as
	41	88.8	1.9	3621	12	ADN71923	Adn71923 Human DKF
C	42	88.4	1.9	455	5	ABV22061	Abv22061 Human pro
C	43	88.4	1.9	455	5	ABV27896	Abv27896 Human pro
C	44	88.4	1.9	476	5	ABV12334	Abv12334 Human pro
C	45	88.4	1.9	504	5	ABV33478	Abv33478 Human pro

ALIGNMENTS

RESULT 1

AAV03326

ID AAV03326 standard; DNA; 4596 BP.

XX AC AAV03326;

XX AC AAV03326;

DT 15-APR-1998 (first entry)

XX DB DNA sequence of NF-kappaB inducing kinase.

XX KW Human tumour necrosis factor receptor-associated factor 2; TRAF2;

XX KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;

XX KW intracellular signalling activity; acute hepatitis;

XX KW autoimmune-induced cell death; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 232..3075

XX FT FT /\*tag= a

XX PN WO9737016-A1.

XX PD 09-OCT-1997.

XX PF 01-APR-1997; 97WO-IL000117.

XX PR 02-APR-1996; 96IL-00117800.

XX PR 26-AUG-1996; 96IL-00119133.

XX XX (YEDA ) YEDA RES & DEV CO LTD.

XX PA Aav71603 Human NF-

XX XX Aax87842 NF-kB ind

XX PI Adq67410 Novel hum

XX XX Acn44437 Mouse mRN

XX DR Acn44437 Mouse mRN

XX DR P-PSDB; AAW42402.

XX PI Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX WPI; 1997-503101/46.

XX DR P-PSDB; AAW42402.

XX FT DNA encoding tumour necrosis factor receptor-associated factor binding

XX PT molecule - used for modulation or mediation in cells of the activity of

XX NF-kB.

XX PS Claim 11; Fig 6; 127pp; English.

XX XX

CC	The present sequence encodes a NF-kappa inducing kinase (NIK). The full length sequence was obtained by PCR using clone 10 (AAV03324). NIK was found to induce NF-kappaB even more effectively than TRAF2. Proteins capable of binding to TRAF2 and NIK were identified. The NIK or TRAF-2 binding proteins can be used for modulation or mediation in cells of NF-kappaB activity or any other intracellular signalling activity modulated or mediated by TRAF2. TRAF-binding proteins are especially used for prevention or treatment of pathological conditions associated with NF-kappaB induction, e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the beta Langerhans cells or the pancreas that results in diabetes, the death of cells in graft rejection, the death of oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited T cell suicide which causes proliferation of the AIDS virus and hence the AIDS disease. The proteins are also useful for screening of ligands capable of binding to a protein, which are useful for modulating cellular activity modulated/mediated by TRAF2
XX	
SQ	Sequence 4596 BP; 1055 A; 1402 C; 1298 G; 841 T; 0 U; 0 Other;
	Query March 100.0%; Score 4596; DB 2; Length 4596; Best Local Similarity 100.0%; Mismatches 0; Matches 4596; Conservative 0; Indels 0; Gaps 0;
QY	1 AGCGGGGGAATGTGCCGTGTGGAAACGTGTAGCTTTTGAAGTGGACTCTGTACCATTG 60
DB	1 AGCGGGGGAATGTGCCGTGTGGAAACGTGTAGCTTTTGAAGTGGACTCTGTACCATTG 60
QY	61 AGGATGTTTGAGGATGAGTATGTGTGCAGAGGCACATAAAGCAGCAGAGACCCCTT 120
DB	61 AGGATGTTTGAGGATGAGTATGTGTGCAGAGGCACATAAAGCAGCAGAGACCCCTT 120
QY	121 GCCCTGCTTTTCTCCCCCAACCCAGCGTAGCTGTCTTCCCAGGTCTGGGATTCCTA 180
DB	121 GCCCTGCTTTTCTCCCCCAACCCAGCGTAGCTGTCTTCCCAGGTCTGGGATTCCTA 180
QY	181 AGTGACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTTGGAGATGCGAGT 240
DB	181 AGTGACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCTTGGAGATGCGAGT 240
QY	241 ATGGAAATGGCTGCGCAGGTGCCCTCGCTCAGCAGTGGGCGACAGAAGAACTCCCC 300
DB	241 ATGGAAATGGCTGCGCAGGTGCCCTCGCTCAGCAGTGGGCGACAGAAGAACTCCCC 300
QY	301 AAGCCAAAAGGAGAACGCGCCACTTGGGGAGAAACAAGAGCTCCGTCTCAAAGCTTGAG 360
DB	301 AAGCCAAAAGGAGAACGCGCCCACTGGGGAGAAACAAGAGCTCCGTCTCAAAGCTTGAG 360
QY	361 GCCGTGGAGAGAGACCCCTGTGTCTTCGGAAAGTGGGAGATCCTGAAGACCTGATTACC 420
DB	361 GCCGTGGAGAGAGACCCCTGTGTCTTCGGAAAGTGGGAGATCCTGAAGACCTGATTACC 420
QY	421 AAGGGCACAGCCAAAGAAAGGCTCCGAGGCGAGGGCCAGTGCCTATCTATCATCGCCAG 480
DB	421 AAGGGCACAGCCAAAGAAAGGCTCCGAGGCGAGGGCCAGTGCCTATCTATCATCGCCAG 480
QY	481 GGTGAGTGTGAGAAATAGCCAAGAGTTACGCCCACTTTTCAGAACGCAATTTTCATCGCT 540
DB	481 GGTGAGTGTGAGAAATAGCCAAGAGTTACGCCCACTTTTCAGAACGCAATTTTCATCGCT 540
QY	541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTTGATCAGATCCCCAAACAAATGTGGCCCAT 600
DB	541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTTGATCAGATCCCCAAACAAATGTGGCCCAT 600
QY	601 GGTACAGAGGGCAAAATAGCCCGTGTGTGTGGAAAGGAAAGCGTTCGCAAGAAAGCCCGG 660
DB	601 GGTACAGAGGGCAAAATAGCCCGTGTGTGTGGAAAGGAAAGCGTTCGCAAGAAAGCCCGG 660
QY	661 AAGAAACGGAAGAAAGAGAGCTCAAGTCCCTGGCTCATGCGAGGTGGCTTGGCCAAA 720
DB	661 AAGAAACGGAAGAAAGAGAGCTCAAGTCCCTGGCTCATGCGAGGTGGCTTGGCCAAA 720
QY	721 CCCCTCCCCAGGACCCCTCAGCAGGAGAGCTGCACCATCCAGTCAGGAGGATGAGTCT 780



Db 4021 CAGAGCAGAGCTCAGCATCACACTGACACTCACCTGCCCCCTGCCCCAGAGGGTAC 4080  
QY |||||  
Db 4081 TGGCAGCGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAG 4140  
Db 4081 TGGCAGCGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAG 4140  
QY 4141 GGCAGGGCAGGCGAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGAC 4200  
Db 4141 GGCAGGGCAGGCGAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGAC 4200  
QY 4201 ACAGTCTTGAGCTGTCCACATGCATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGA 4260  
Db 4201 ACAGTCTTGAGCTGTCCACATGCATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGA 4260  
QY 4261 ATAGACACCCCTTCCCATTCGCCAGCTTAGCCTCTTCCAGGGGAGCTACTCAGGA 4320  
Db 4261 ATAGACACCCCTTCCCATTCGCCAGCTTAGCCTCTTCCAGGGGAGCTACTCAGGA 4320  
QY 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAACTCTCT 4380  
Db 4321 CTCAGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAACTCTCT 4380  
QY 4381 GGGCAGGGGACGCCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCACGACAGC 4440  
Db 4381 GGGCAGGGGACGCCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCACGACAGC 4440  
QY 4441 CTTTGTCCAGCTGTCCACATTCAGTCAGCTCTCCGGGGAGAGAGCCCGGCCCCAG 4500  
Db 4441 CTTTGTCCAGCTGTCCACATTCAGTCAGCTCTCCGGGGAGAGAGCCCGGCCCCAG 4500  
QY 4501 CACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAG 4560  
Db 4501 CACATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAG 4560  
QY 4561 CAATAAAGTTGGGGTGTAGCAAAATGTTAAAAAAA 4596  
Db 4561 CAATAAAGTTGGGGTGTAGCAAAATGTTAAAAAAA 4596

RESULT 2  
AAS79490  
ID AAS79490 standard; cDNA; 4596 BP.

XX AAS79490;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #15294.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG15303.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX Claim 1; SEQ ID NO 15294; 103pp; English.  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;

Query Match 99.6%; Score 4578.4; DB 5; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 5 GGGGACTGTGCGGTGTGGAACGTGTAGCTTTGA-AGTGGACTCTGTACCATTGAGG 63  
Db 5 GGGGACTGTGCGGTGTGGAACGTGTAGCTTTGAAGGTGGACTCTGTACCATTGAGG 64  
QY 64 ATGTTTGAGGATGATGTGTGCGAGGACACATATAACAGCAGACACCTTTTGGC 123  
Db 65 ATGTTTGAGGATGATGTGTGCGAGGACACATATAACAGCAGACACCTTTTGGC 124  
QY 124 COTGCTTTTCCCCCAACCCAGGCTGACCTGTGTTCTTCCCAGGCTTGGGATTTAAAGT 183  
Db 125 COTGCTTTTCCCCCAACCCAGGCTGACCTGTGTTCTTCCCAGGCTTGGGATTTAAAGT 184  
QY 184 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAGCTGGGAGATGGCAGTGATG 243  
Db 185 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAGCTGGGAGATGGCAGTGATG 244  
QY 244 GAAATGGGCTGCCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCCAAG 303  
Db 245 GAAATGGGCTGCCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCCAAG 304  
QY 304 CCAAAAGGAGAGACGCCGCCACTGCTGGGGAAGAAACAGAGCTCCGTTACAAAGTTGAGGCC 363  
Db 305 CCAAAAGGAGAGACGCCGCCACTGCTGGGGAAGAAACAGAGCTCCGTTACAAAGTTGAGGCC 364  
QY 364 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTGAATGACGTGATTACCAAG 423  
Db 365 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCTGAATGACGTGATTACCAAG 424  
QY 424 GGCACAGCCCAAGGAGGCTCCGAGGCGAGCTGCATCTCTATCATGCCCCAGGCT 483  
Db 425 GGCACAGCCCAAGGAGGCTCCGAGGCGAGCTGCATCTCTATCATGCCCCAGGCT 484  
QY 484 GAGTGTGAGAAATAGCCAAAGAGTTTACAGCCCCACCTTTTTCAGAACGCAATTTTCATCGTGGG 543  
Db 485 GAGTGTGAGAAATAGCCAAAGAGTTTACAGCCCCACCTTTTTCAGAACGCAATTTTCATCGTGGG 544  
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGCCCCATGCT 603  
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGCCCCATGCT 604

Qy	604	ACAGAGGCAAAATGGCCCGCTGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	663
Db	605	ACAGAGGCAAAATGGCCCGCTGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG	664
Qy	664	AAACGGAAGAAAGAGAGCTCAAGTCCCTGCTCATGACGAGGTGGCTTGGCCAAACCC	723
Db	665	AAACGGAAGAAAGAGAGCTCAAGTCCCTGCTCATGACGAGGTGGCTTGGCCAAACCC	724
Qy	724	CTCCCCAGGACCCCTGACGAGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCCA	783
Db	725	CTCCCCAGGACCCCTGACGAGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCCA	784
Qy	784	CTCGCGCCCATATGTTAGAAAACCCCGCAGTTACCAAGCCTCTGAAGAAACAGGCG	843
Db	785	CTCGCGCCCATATGTTAGAAAACCCCGCAGTTACCAAGCCTCTGAAGAAACAGGCG	844
Qy	844	CTTGGCAACTCTGTTTTAAGCAGCTTGGCAGAGGCTTACGCGCGCTCTCCCTCGATCA	903
Db	845	CTTGGCAACTCTGTTTTAAGCAGCTTGGCAGAGGCTTACGCGCGCTCTCCCTCGATCA	904
Qy	904	GAACTCCACAAACTGATCAGCGCCCTTGCATGTCTGAACCACTGTGGAACCTGCACCCAC	963
Db	905	GAACTCCACAAACTGATCAGCGCCCTTGCATGTCTGAACCACTGTGGAACCTGCACCCAC	964
Qy	964	CCCCAGGAGGAGGCGCCCTGCGCCCTGCCCCAGCACCCCTTCCCTATAGCAGACTGCT	1023
Db	965	CCCCAGGAGGAGGCGCCCTGCGCCCTGCCCCAGCACCCCTTCCCTATAGCAGACTGCT	1024
Qy	1024	CATCCCTTCCCATTTCCACCTCTCAGCCCTGGAACCTCACTCTGGAGTCTTCCCTG	1083
Db	1025	CATCCCTTCCCATTTCCACCTCTCAGCCCTGGAACCTCACTCTGGAGTCTTCCCTG	1084
Qy	1084	GGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACACTGAGCAAACTG	1143
Db	1085	GGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACACTGAGCAAACTG	1144
Qy	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCCACACTGAGAGCCAGTCTGTCT	1203
Db	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCCACACTGAGAGCCAGTCTGTCT	1204
Qy	1204	CGTGGTCCCATGAGAGTTTCTGTGGAGAAATACCTAGTGCATGTCTGCAAGGCGAGC	1263
Db	1205	CGTGGTCCCATGAGAGTTTCTGTGGAGAAATACCTAGTGCATGTCTGCAAGGCGAGC	1264
Qy	1264	GTGAGCTCAAGCCAGGCCCCAGCTGACAGCTTGGCCAGAGCTGGGAGCAGCGGGC	1323
Db	1265	GTGAGCTCAAGCCAGGCCCCAGCTGACAGCTTGGCCAGAGCTGGGAGCAGCGGGC	1324
Qy	1324	TCCAGATCCCGGAGCCAGCCCCAAACTCAGGACAAACGAGGTGTCTGCTCACTGAG	1383
Db	1325	TCCAGATCCCGGAGCCAGCCCCAAACTCAGGACAAACGAGGTGTCTGCTCACTGAG	1384
Qy	1384	AAACTCAAGCCAGTGGATTATGATACCGAAGAGTCCACTGGGCCAGCCACAGCTC	1443
Db	1385	AAACTCAAGCCAGTGGATTATGATACCGAAGAGTCCACTGGGCCAGCCACAGCTC	1444
Qy	1444	CGCTTGGCAGAGGCTCTCTCGAGAGGTGCACAGGATGGAGGACAGCACTGGCTTC	1503
Db	1445	CGCTTGGCAGAGGCTCTCTCGAGAGGTGCACAGGATGGAGGACAGCACTGGCTTC	1504
Qy	1504	CAGTGGCTGTCAAAAGGTGGCTTGAAGTATTTGGGCGAGAGGCTGATGGCATGT	1563
Db	1505	CAGTGGCTGTCAAAAGGTGGCTTGAAGTATTTGGGCGAGAGGCTGATGGCATGT	1564
Qy	1564	GCAGGATTGACTCACCCAGAAATGTCTCTTGTATGGAGCTGTGAGAAAGGCTTGG	1623
Db	1565	GCAGGATTGACTCACCCAGAAATGTCTCTTGTATGGAGCTGTGAGAAAGGCTTGG	1624
Qy	1624	GTCAACATCTTTCATGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGAGTGGTCAAGAGCAG	1683
Db	1625	GTCAACATCTTTCATGGAGCTCTGGAAGGTGGCTCCCTGGGCCAGAGTGGTCAAGAGCAG	1684
Qy	1684	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACTTGGGCCAGGCGCTGGAGGGTCTGGAA	1743

Db	1685	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACTTGGGCCAGGCGCTTGGAAGGGTCTGGAA	1744
Qy	1744	TACCTCCACTCACAAGGATTTCTCATGGGAGCTCAAAAGCTGACAAAGCTCTCTGTGCC	1803
Db	1745	TACCTCCACTCACAAGGATTTCTCATGGGAGCTCAAAAGCTGACAAAGCTCTCTGTGCC	1804
Qy	1804	AGCATGGAGCCACGAGGCCCTCTGTGACTTTGGCCATGTGTGTCTTCAACCTGAT	1863
Db	1805	AGCATGGAGCCACGAGGCCCTCTGTGACTTTGGCCATGTGTGTCTTCAACCTGAT	1864
Qy	1864	GGCTTGGGAAAGTCTTCTCACAAGGGACTACATCCCTTGGCAGAGAGCCACATGGCT	1923
Db	1865	GGCTTGGGAAAGTCTTCTCACAAGGGACTACATCCCTTGGCAGAGAGCCACATGGCT	1924
Qy	1924	CCGAGGTGGTGTGGGCGAGGCTGCGAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1983
Db	1925	CCGAGGTGGTGTGGGCGAGGCTGCGAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1984
Qy	1984	ATGATGTGCACATGTCTCAAGCGCTGCCACCCCTGGACTCAAGTTCTTCCGAGGGCGCTC	2043
Db	1985	ATGATGTGCACATGTCTCAAGCGCTGCCACCCCTGGACTCAAGTTCTTCCGAGGGCGCTC	2044
Qy	2044	TGCTCAAGATTGCCAGGAGCTTCCGCTGTGGAGGATCCCACTCTCTGGCCCCCT	2103
Db	2045	TGCTCAAGATTGCCAGGAGCTTCCGCTGTGGAGGATCCCACTCTCTGGCCCCCT	2104
Qy	2104	CTCAGCCCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2163
Db	2105	CTCAGCCCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2164
Qy	2164	GGGAGCTGGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGGCCCT	2223
Db	2165	GGGAGCTGGGAGGAAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAAGGCCCT	2224
Qy	2224	TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATTAACACAG	2283
Db	2225	TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAATCAAGCCAAATTAACACAG	2284
Qy	2284	ACCTTCATGCCAGCGAGAGCTTTCGCAAGGGCCCCAGGGCCCCCGGCCAGCTGAG	2343
Db	2285	ACCTTCATGCCAGCGAGAGCTTTCGCAAGGGCCCCAGGGCCCCCGGCCAGCTGAG	2344
Qy	2344	GAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCCACAGAGCCA	2403
Db	2345	GAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCCACAGAGCCA	2404
Qy	2404	AACAAGTCTCTCTCTTGAATTTTGAGCAAGAGAGTCTGGGATGTGGGAAACCTTACCT	2463
Db	2405	AACAAGTCTCTCTCTTGAATTTTGAGCAAGAGAGTCTGGGATGTGGGAAACCTTACCT	2464
Qy	2464	CTGTCTCTCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465	CTGTCTCTCTGGAGCCAGCCCTGCGCAAGAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524	GTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAAATTTCTTCAACAGCCTGTCCAG	2583
Db	2525	GTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAAATTTCTTCAACAGCCTGTCCAG	2584
Qy	2584	CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTGTGTCTCTCAGCATCGACAGCTCTCC	2643
Db	2585	CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTGTGTCTCTCAGCATCGACAGCTCTCC	2644
Qy	2644	CTGTCCGATCAGAGTGAAGAACCCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG	2703
Db	2645	CTGTCCGATCAGAGTGAAGAACCCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG	2704
Qy	2704	AGCTCAGCGCTACACTCTCTGGAGCAGCGAGGCTCGAAGCTCCAGCTGTGAAACATG	2763
Db	2705	AGCTCAGCGCTACACTCTCTGGAGCAGCGAGGCTCGAAGCTCCAGCTGTGAAACATG	2764
Qy	2764	GTGTGCGCCCGGGCGGCCACCGCACCCCAAGCTATTTCAATGTGTGAAAGTCCAA	2823



Db	2765	GTGTTGCCCGGGGGGGCCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAATGCCAA	2824
Qy	2824	ATACAGTCTCTTAATGGTGAACACCTTGCAATCCGGGAGTTCCACGGGTCAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGGTGAACACCTTGCAATCCGGGAGTTCCACGGGTCAAAGTGGGA	2884
Qy	2884	GACATCCCACTGGCATCAGCAGCAGATCCCAAGCTTCAGCTTCAGTTGGTTCACCAAA	2943
Db	2885	GACATCCCACTGGCATCAGCAGCAGATCCCAAGCTTCAGCTTCAGTTGGTTCACCAAA	2944
Qy	2944	GACGGGAGCCTGTTTCCTACGACATGGAGGTGCCAGACTTCGGGCAATCGACCTGCAGTGC	3003
Db	2945	GACGGGAGCCTGTTTCCTACGACATGGAGGTGCCAGACTTCGGGCAATCGACCTGCAGTGC	3004
Qy	3004	ACACTGCCCTTGATGTCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCATGGCAGCTGGAG	3063
Db	3005	ACACTGCCCTTGATGTCAGCTTCGCCCTGGAGCTGGAGGGTCAAGCATGGCAGCTGGAG	3064
Qy	3064	AACAGGGCCCTAACCCCTGCCCTCCACGCGCGGCTCCACACTGCCGGAAGACAGCCTTCCTG	3123
Db	3065	AACAGGGCCCTAACCCCTGCCCTCCACGCGCGGCTCCACACTGCCGGAAGACAGCCTTCCTG	3124
Qy	3124	CTCGGTGCACGATGCTGCTCCCTGAAAAACAAGGCTCAGCCGTTCCCAAGGGATTCGACGC	3183
Db	3125	CTCGGTGCACGATGCTGCTCCCTGAAAAACAAGGCTCAGCCGTTCCCAAGGGATTCGACGC	3184
Qy	3184	CCCGGCTCAGAGTGGGAAACGAGGCTCCGACGACCAAGGTCGGGGCAGCAGATGCC	3243
Db	3185	CCCGGCTCAGAGTGGGAAACGAGGCTCCGACGACCAAGGTCGGGGCAGCAGATGCC	3244
Qy	3244	TCCAGGATTTTCAACACTGAGCCCTGCCCCCACCCTGCTGAAAAACAATCCGCCACGTGAA	3303
Db	3245	TCCAGGATTTTCAACACTGAGCCCTGCCCCCACCCTGCTGAAAAACAATCCGCCACGTGAA	3304
Qy	3304	GAGACAGAAGGAGATGGCAGGATTTACCTGGGGAAACAAACAAGGATCTTTTTCTGCC	3363
Db	3305	GAGACAGAAGGAGATGGCAGGATTTACCTGGGGAAACAAACAAGGATCTTTTTCTGCC	3364
Qy	3364	CCTGCTCCAGTTCGAGTTGGGCTGACCCGCTTGGATCAGTACCATTTGTTGGCAGACAGG	3423
Db	3365	CCTGCTCCAGTTCGAGTTGGGCTGACCCGCTTGGATCAGTACCATTTGTTGGCAGACAGG	3424
Qy	3424	GGAGAGCAGTTCACAGCTTGGGTGAGAGGGGTGGGAGCCCTTCGCGCCCTCACCCCTC	3483
Db	3425	GGAGAGCAGTTCACAGCTTGGGTGAGAGGGGTGGGAGCCCTTCGCGCCCTCACCCCTC	3484
Qy	3484	CAGGCTGCTGTCAGAGTCTCAAGTGTGTAAAGGGCCAAACTCAGGTTTCAGTCAGAACCA	3543
Db	3485	CAGGCTGCTGTCAGAGTCTCAAGTGTGTAAAGGGCCAAACTCAGGTTTCAGTCAGAACCA	3544
Qy	3544	GGTCAGCAGGTATGCCCGCGGTAGGTTTAAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC	3603
Db	3545	GGTCAGCAGGTATGCCCGCGGTAGGTTTAAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC	3604
Qy	3604	ACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGATGCTGACCTGGGAGGCT	3663
Db	3605	ACCTGGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGATGCTGACCTGGGAGGCT	3664
Qy	3664	CCCTGGTAGATACACACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTCAACCTCT	3723
Db	3665	CCCTGGTAGATACACACACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTCAACCTCT	3724
Qy	3724	GGTTCAGCCAAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCAGACATGT	3783
Db	3725	GGTTCAGCCAAGGACCAAGAGGTGTGTAAAGTGAAGTGGTTCTCAGTCCCAGACATGT	3784
Qy	3784	GGCCCTTGTGCTGGCTACCACTCTCCACAGAGCAGCGCCCCCGAGCCCTTCAGGC	3843
Db	3785	GGCCCTTGTGCTGGCTACCACTCTTCCACAGACAGCGCCCCCGAGCCCTTCAGGC	3844
Qy	3844	CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTAT	3903
Db	3845	CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTAT	3904

QY	3904	GCAGGATATATGCCCTGCACAGGAACAGTCTGTGGATGGACATGATCAGTGCCTTAAGGAAGCAG	3964
DB	3905	GCAGGATATATGCCCTGCACAGGAACAGTCTGTGGATGGACATGATCAGTGCCTTAAGGAAGCAG	3964
QY	3964	CAGAGAGAGACGTCGCGGCCCCAGCCCCACTATCAGTGTCCAGCGTGTCTGGTTCGCCAG	4023
DB	3965	CAGAGAGAGACGTCGCGGCCCCAGCCCCACTATCAGTGTCCAGCGTGTCTGGTTCGCCAG	4024
QY	4024	AGCACAGCTCAGCATCACTGAACCTCAACCTTGCCTTGCCTTGGCCAGAGGGTACTGC	4083
DB	4025	AGCACAGCTCAGCATCACTGAACCTCAACCTTGCCTTGCCTTGGCCAGAGGGTACTGC	4084
QY	4084	CGACGGCACTTTGGCACTCTGATGACCTCAAGACACTTTTCATGCTGCCCTCTGCGAGGGC	4143
DB	4085	CGACGGCACTTTGGCACTCTGATGACCTCAAGACACTTTTCATGCTGCCCTCTGCGAGGGC	4144
QY	4144	AGGCGAGGAGTGCACACTGTAGGAGCATATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4203
DB	4145	AGGCGAGGAGTGCACACTGTAGGAGCATATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4204
QY	4204	GTCTTGAGCTGTCCACATGCATGTGAATCTCTCAAACTCTTCCAGATTTCTCTAAGAATA	4263
DB	4205	GTCTTGAGCTGTCCACATGCATGTGAATCTCTCAAACTCTTCCAGATTTCTCTAAGAATA	4264
QY	4264	GCACCCCTTCCCATTTGCCCCAGCTTAGCTCTTCTCCAGGGAGCTACTCAGGACTC	4323
DB	4265	GCACCCCTTCCCATTTGCCCCAGCTTAGCTCTTCTCCAGGGAGCTACTCAGGACTC	4324
QY	4324	ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCTGGG	4383
DB	4325	ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCTGGG	4384
QY	4384	GCAGGGAGCGCCGAGACTCCTCGTGGAGAGCTCAATCCACACTTTCGCCAAGACAGCCTT	4443
DB	4385	GCAGGGAGCGCCGAGACTCCTCGTGGAGAGCTCAATCCACACTTTCGCCAAGACAGCCTT	4444
QY	4444	TGTCAGCTGTCCACATTTGAGTCAGACTGTCTCCCGGGAGAGAGCCCGGCCCCAGCAC	4503
DB	4445	TGTCAGCTGTCCACATTTGAGTCAGACTGTCTCCCGGGAGAGAGCCCGGCCCCAGCAC	4504
QY	4504	ATAAGAACTCGAGCCTTTGGTACTGCAGAGTCTGGGTGTAGAGAACTCTTTTGAAGCAA	4563
DB	4505	ATAAGAACTCGAGCCTTTGGTACTGCAGAGTCTGGGTGTAGAGAACTCTTTTGAAGCAA	4564
QY	4564	TAAAGTTTGGGTGTATGACAAATGTTAAAAA	4595
DB	4565	TAAAGTTTGGGTGTATGACAAATGTTAAAAA	4596
RESULT 3			
ABKS1171			
ID	ABKS1171 standard; cDNA; 4596 BP.		
XX	XX		
AC	ABKS1171;		
XX	XX		
DT	30-JUL-2002 (first entry)		
XX	XX		
DE	cDNA encoding human cellular kinase NIK protein.		
KW	Human; viricide; cytomegalovirus infection; CMV; cellular kinase; RICK;		
KW	RIP; Nck-Interacting kinase; NIK; MKK3; SRPK-2; gene; ss.		
XX	Homo sapiens.		
XX	XX		
Key	Location/Qualifiers		
FT	233..3076		
CDS	/*tag= a		
FT	/product= "Human cellular kinase NIK (Nck-Interacting		
FT	Kinase)"		
XX	XX		
PN	EP1201765-A2.		
XX	XX		



PD 02-MAY-2002.  
XX PF 15-OCT-2001; 2001EP-00124604.  
XX PR 16-OCT-2000; 2000US-0240750P.  
XX PA (AXXI-) AXIWA PHARM AG.  
XX Schubart D, Habenberger P, Stein-Gerlach M, Bevec D;  
XX WPI; 2002-373930/41.  
XX DR P-PSDB; AAU80371.  
XX  
XX Identifying agents for treatment or prevention of cytomegalovirus  
PT infection, comprises contacting test compound with cellular kinase and  
PT detecting change in cellular kinase activity.  
XX  
XX Disclosure; Page 29-34; 49pp; English.  
XX  
XX The present invention relates to a new method for identifying compounds  
CC for treating and/or preventing cytomegalovirus (CMV) infection and/or  
CC related diseases. The method of the invention comprises contacting a test  
CC compound with at least one of the cellular kinases RICK, RIP, Nck,  
CC Interacting kinase, MKK3 and SRPK-2 and detecting any change in kinase  
CC activity. The method of the invention can be used to treat and/or prevent  
CC CMV infections and related diseases. Oligonucleotides that can detect the  
CC specified kinases can also be used for diagnosis of infection. The  
CC present nucleic acid sequence encodes the human cellular kinase NIK (Nck-  
CC Interacting kinase) protein of the invention, as described above  
XX  
XX Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;  
XX  
XX Query Match 99.6%; Score 4578.4; DB 6; Length 4596;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
XX  
Qy 5 GGGGAGCTGTCCGTGTGGAACGTGTAGCTGTGAGAGGTGGAATCTGTTACCAATTGAGG 63  
Db 5 GGGGAGCTGTCCGTGTGGAACGTGTAGCTGTGAGAGGTGGAATCTGTTACCAATTGAGG 64  
Qy 64 ATGTTTGGAGATGAGTATGTGTGGCAGAGGACACATTAACAGCGAGACCTTTTGCC 123  
Db 65 ATGTTTGGAGATGAGTATGTGTGGCAGAGGACACATTAACAGCGAGACCTTTTGCC 124  
Qy 124 CCTGCTTTCCTCCCAACCAAGCTGACCTGTGTCTCCAGGTCTGGGATCTTAAGT 183  
Db 125 CCTGCTTTCCTCCCAACCAAGCTGACCTGTGTCTCCAGGTCTGGGATCTTAAGT 184  
Qy 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCCTGGGAGTGGCAGTGATG 243  
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCCTGGGAGTGGCAGTGATG 244  
Qy 244 GAAATGGCTCCCAAGTGGCCCTGCTCAGCAGTGGGCGAGAGAGAACTCCCAAG 303  
Db 245 GAAATGGCTCCCAAGTGGCCCTGCTCAGCAGTGGGCGAGAGAGAACTCCCAAG 304  
Qy 304 CCAAGGAGAGAGACGCGCCACTCGGGAAGAAACAGAGCTCCGTCTCAAGCTTTGAGGCC 363  
Db 305 CCAAGGAGAGAGACGCGCCACTCGGGAAGAAACAGAGCTCCGTCTCAAGCTTTGAGGCC 364  
Qy 364 GTGGAGAAGAGCCCTGTGTTCTCGGGAAGTGGGAGATCCTGAATGACGTGATTAACCAAG 423  
Db 365 GTGGAGAAGAGCCCTGTGTTCTCGGGAAGTGGGAGATCCTGAATGACGTGATTAACCAAG 424  
Qy 424 GGCAACAGCAAGAGGCTCCGAGGAGGGCCAGCTGCGCATCTTATCATCGCCAGGCT 483  
Db 425 GGCAACAGCAAGAGGCTCCGAGGAGGGCCAGCTGCGCATCTTATCATCGCCAGGCT 484  
Qy 484 GAGTGTGAGATAGCAAGAGTTCAGGCCACCTTTTCAGAACCGCATTTTCATCGCTGGG 543  
Db 485 GAGTGTGAGATAGCAAGAGTTCAGGCCACCTTTTCAGAACCGCATTTTCATCGCTGGG 544  
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 603  
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 604  
Qy 604 ACAGAGGGCAAAATGGCCCGTGTGTTGGAAGGAAAGCGTTCGACGAAAGCCGGGAG 663  
Db 605 ACAGAGGGCAAAATGGCCCGTGTGTTGGAAGGAAAGCGTTCGACGAAAGCCGGGAG 664  
Qy 664 AAACGGAAGAAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCTTTGGCCAAACCC 723  
Db 665 AAACGGAAGAAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCTTTGGCCAAACCC 724  
Qy 724 CTCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCTCCA 783  
Db 725 CTCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCTCCA 784  
Qy 784 CTCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCTCCA 843  
Db 785 CTCCGAGGACCCCTGAGCAGGAGAGCTGCACCATCCAGTGCAGGAGGATGAGTCTCCA 844  
Qy 844 CTTCGGCAACTCTGTTTAAAGCAGCTTTGGGAGGGCTACGGCCGGCTCTGCTCGATCA 903  
Db 845 CTTCGGCAACTCTGTTTAAAGCAGCTTTGGGAGGGCTACGGCCGGCTCTGCTCGATCA 904  
Qy 904 GAACTCCAAACATGATCAGCCCTTGCATATGTTGAACACACGTTGTGGAACCTGCACAC 963  
Db 905 GAACTCCAAACATGATCAGCCCTTGCATATGTTGAACACACGTTGTGGAACCTGCACAC 964  
Qy 964 CCCGAGGAGGAGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTATAGCAGACTGCT 1023  
Db 965 CCCGAGGAGGAGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTATAGCAGACTGCT 1024  
Qy 1024 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCCCTCTGGAGTCTCTCTG 1083  
Db 1025 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCCCTCTGGAGTCTCTCTG 1084  
Qy 1084 GGCATACTGCGCTGTGTAGACAGCCAGAAACCTTGTGCTGACCCACACCTGTAGCAAACTG 1143  
Db 1085 GGCATACTGCGCTGTGTAGACAGCCAGAAACCTTGTGCTGACCCACACCTGTAGCAAACTG 1144  
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCGCCAGCAGCCCTGGAAGCCAGCTGCTGTCT 1203  
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCGCCAGCAGCCCTGGAAGCCAGCTGCTGTCT 1204  
Qy 1204 CGTGTGTCATGAGAAAGTCTGTGTGAGGAAATACCTAGTGTGATGCTCTGCAAGGAGC 1263  
Db 1205 CGTGTGTCATGAGAAAGTCTGTGTGAGGAAATACCTAGTGTGATGCTCTGCAAGGAGC 1264  
Qy 1264 GTGAGCTCAAGCCAGGCGCCACAGCTGACCTGGCCAAAGACCTGGGAGCAGCAGGGGC 1323  
Db 1265 GTGAGCTCAAGCCAGGCGCCACAGCTGACCTGGCCAAAGACCTGGGAGCAGCAGGGGC 1324  
Qy 1324 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACCAACGAGGGTGTCTGCTCACTGAG 1383  
Db 1325 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACCAACGAGGGTGTCTGCTCACTGAG 1384  
Qy 1384 AAATCAAGCAGTGGATATGATACCGGAGAAAGTCCACTGGGGCCACGACCAAGCTC 1443  
Db 1385 AAATCAAGCAGTGGATATGATACCGGAGAAAGTCCACTGGGGCCACGACCAAGCTC 1444  
Qy 1444 CGCTTGGGAGAGGCTCTTTCGGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1503  
Db 1445 CGCTTGGGAGAGGCTCTTTCGGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1504  
Qy 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGGAGCTGATGGCATGT 1563  
Db 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGGAGCTGATGGCATGT 1564  
Qy 1564 GCAGGATTTGACCTCACCCAGAAATTTGCTCCCTTGTATGAGAGTGTGAGAGAGGGCTTGG 1623  
Db 1565 GCAGGATTTGACCTCACCCAGAAATTTGCTCCCTTGTATGAGAGTGTGAGAGAGGGCTTGG 1624  
Qy 1624 GTCAACATCTTATGAGAGTGTGGAGGTGGCTCCCTGGGCGCAGCTGGTCAAGGAGCAG 1683

Db 1625 GTCAACATCTTCATGGAGTGCTGGAAGGTGGCTCCCTGGGCCAGTGGTCAAGAGCAG 1684  
QY 1684 GGCTGTCTCCAGAGACCGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGTCGAA 1743  
Db 1685 GGCTGTCTCCAGAGAACGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGTCGAA 1744  
QY 1744 TACCTCCATCTACGAAGGATCTGCATGGGACGTCGAAAGCTGACAAAGTCGTCTCTGTCC 1803  
Db 1745 TACCTCCATCTACGAAGGATCTGCATGGGACGTCGAAAGCTGACAAAGTCGTCTCTGTCC 1804  
QY 1804 AGCGATGGAGCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGAT 1863  
Db 1805 AGCGATGGAGCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGAT 1864  
QY 1864 GSCCTGGGAAAGTCTTGCTCAGCGGACTACATCCCTGGGCACAGAGACCCACATGGCT 1923  
Db 1865 GSCCTGGGAAAGTCTTGCTCAGCGGAGACTACATCCCTGGGCACAGAGACCCACATGGCT 1924  
QY 1924 CCGGAGGTGGTGTCTGGGACGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGT 1983  
Db 1925 CCGGAGGTGGTGTCTGGGACGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGT 1984  
QY 1984 ATGATGCTGCACATGCTCAAGCGGCTGCACCCCTGGACTCAGTTCTTCCGAGGGCGGCTC 2043  
Db 1985 ATGATGCTGCACATGCTCAAGCGGCTGCACCCCTGGACTCAGTTCTTCCGAGGGCGGCTC 2044  
QY 2044 TSCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTGGCCCT 2103  
Db 2045 TSCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTGGCCCT 2104  
QY 2104 CTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCACCGCGTGTCTGCA 2163  
Db 2105 CTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCACCGCGTGTCTGCA 2164  
QY 2164 GCGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCGAAAGACCCCT 2223  
Db 2165 GCGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGTCTGAAAGACCCCT 2224  
QY 2224 TGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCACAG 2283  
Db 2225 TGGAGGGAGATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCACAG 2284  
QY 2284 ACCCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCAAGGGCCCGGCCAGCTGAG 2343  
Db 2285 ACCCTCCATGCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCAAGGGCCCGGCCAGCTGAG 2344  
QY 2344 GAGACAAAGGAGAGCCCTTAAGCTCAGCCTCTCTCCACAGAGGCCGCCAGAGCA 2403  
Db 2345 GAGACAAAGGAGAGCCCTTAAGCTCAGCCTCTCTCCACAGAGGCCGCCAGAGCA 2404  
QY 2404 AACAACTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACT 2463  
Db 2405 AACAACTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCTTACT 2464  
QY 2464 CTGTCTCTCCAGAGCCCTCTCCAGAAACCCAGCTCACCAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCTCTCCAGAGCCCTCTCCAGAAACCCAGCTCACCAGAGCGGAAAGCAACC 2524  
QY 2524 GTCCGGAGCAGGAAGTCTGAGAGTGGAAATAGAAATTAATCTCAACAGCCTGTCCAG 2583  
Db 2525 GTCCGGAGCAGGAAGTCTGAGAGTGGAAATAGAAATTAATCTCAACAGCCTGTCCAG 2584  
QY 2584 CCATTTTCTCTGGAGAGCAGAGCAATTTCTCTGTGCTCAGCATCGACAGCCTCTCC 2643  
Db 2585 CCATTTTCTCTGGAGAGCAGAGCAATTTCTCTGTGCTCAGCATCGACAGCCTCTCC 2644  
QY 2644 CTGTCTGATGACAGTACAGAAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACACCTG 2703  
Db 2645 CTGTCTGATGACAGTACAGAAACCCATCAAAAGGCTCTCAAAAGCTCGGGGACACCTG 2704  
QY 2704 AGCTCAGGCGTACATCTCTGGAGCAGCCAGGCCAGGCTCGAAGTCTCAGCTGGAAATG 2763  
Db 2705 AGCTCAGGCGTACATCTCTGGAGCAGCCAGGCCAGGCTCGAAGTCTCAGCTGGAAATG 2764

QY 2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAAGTCCAA 2823  
Db 2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAAGTCCAA 2824  
QY 2824 ATACAGTCTCTTAATATGGTGAACACCTGCACATCCGGGAGTTCACCCGGGTCAAGTGGGA 2883  
Db 2825 ATACAGTCTCTTAATATGGTGAACACCTGCACATCCGGGAGTTCACCCGGGTCAAGTGGGA 2884  
QY 2884 GACATCGGCACCTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTTCACCAA 2943  
Db 2885 GACATCGGCACCTGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTTCACCAA 2944  
QY 2944 GACGGCAGCCTGTTGCTACGATGAGGTGCCAGATCGGGCATCGAATTCAGTGC 3003  
Db 2945 GACGGCAGCCTGTTGCTACGATGAGGTGCCAGATCGGGCATCGAATTCAGTGC 3004  
QY 3004 ACATGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGCTGAG 3063  
Db 3005 ACATGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCCAGCTGAG 3064  
QY 3064 AACAGGCCCTAACCTGCCCTCCACCGCGCTCCACACTCCGGAAGCAGCCTTCCTG 3123  
Db 3065 AACAGGCCCTAACCTGCCCTCCACCGCGCTCCACACTCCGGAAGCAGCCTTCCTG 3124  
QY 3124 CTCGCTGACGATGCTGCCCTGAAAAACACAGGCTCAGCCGTTCCAGGGGATTTGCCAGCC 3183  
Db 3125 CTCGCTGACGATGCTGCCCTGAAAAACACAGGCTCAGCCGTTCCAGGGGATTTGCCAGCC 3184  
QY 3184 CCGCGCTCAAGTGGGAAACAGGGCTCGCAGCAGCAAGTGGGGGCAAGCAATGCC 3243  
Db 3185 CCGCGCTCAAGTGGGAAACAGGGCTCGCAGCAGCAAGTGGGGGCAAGCAATGCC 3244  
QY 3244 TCCAGGATTTTCACTGAGCCCTGCCACCTGCTGAAACAAACATCCGCCAGTGAA 3303  
Db 3245 TCCAGGATTTTCACTGAGCCCTGCCACCTGCTGAAACAAACATCCGCCAGTGAA 3304  
QY 3304 GAGCAGAAAGAGATGGCAGAGTTACTGGGAAAAACAAACAGGGATCTTTTCTGCC 3363  
Db 3305 GAGCAGAAAGAGATGGCAGAGTTACTGGGAAAAACAAACAGGGATCTTTTCTGCC 3364  
QY 3364 CTGCTCAGTTCAGTTCGCTGAGCCCTGAGTTCAGTGCATTCATTTGTCGACAGG 3423  
Db 3365 CTGCTCAGTTCAGTTCGCTGAGCCCTGAGTTCAGTGCATTCATTTGTCGACAGG 3424  
QY 3424 GGAGCAGCTTCAGCCTGGGTGAGAGGCTGGGAGCCCTTCGCGCCCTCACCCTC 3483  
Db 3425 GGAGCAGCTTCAGCCTGGGTGAGAGGCTGGGAGCCCTTCGCGCCCTCACCCTC 3484  
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGTTCAGTGCAGAACCA 3543  
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGTTCAGTGCAGAACCA 3544  
QY 3544 GGTTCAGAGGTATGCCCCCGTGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3603  
Db 3545 GGTTCAGAGGTATGCCCCCGTGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3604  
QY 3604 ACCTGGCCAGCTCACCCTTTTGGGTGTAAGGGGAAAGAAATCCCTGAGCCCTGGGAAGCT 3663  
Db 3605 ACCTGGCCAGCTCACCCTTTTGGGTGTAAGGGGAAAGAAATCCCTGAGCCCTGGGAAGCT 3664  
QY 3664 CCCTGGTGAATATACACACATCTTTTCAGTGTGTGCAACACAGGTTCCTGAGTTCACCTCT 3723  
Db 3665 CCCTGGTGAATATACACACATCTTTTCAGTGTGTGCAACACAGGTTCCTGAGTTCACCTCT 3724  
QY 3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTTCGCCAGCATGT 3783  
Db 3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGTCTCAGTTCGCCAGCATGT 3784  
QY 3784 GCCCTTTGCTGCTGCTACCACTTTCCAGAGCAGCAGCCCGGAGCCCTTCAGG 3843  
Db 3785 GCCCTTTGCTGCTGCTACCACTTTCCAGAGCAGCAGCCCGGAGCCCTTCAGG 3844



Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAAAATGTGGCCCATGCT 604  
Qy 604 ACAGAGGGCAAAATGCGCCGTGTGTGTGGAGAGGAAGCGTCGACGAAAGCCCGGAAG 663  
Db 605 ACAGAGGGCAAAATGCGCCGTGTGTGTGGAGAGGAAGCGTCGACGAAAGCCCGGAAG 664  
Qy 664 AAACGGAGAGAGAGAGTCAAAGTCCCTGGCTCATGACGAGAGTGGCCCTTGCCCAAAACC 723  
Db 665 AAACGGAGAGAGAGAGTCAAAGTCCCTGGCTCATGACGAGAGTGGCCCTTGCCCAAAACC 724  
Qy 724 CTCGCCAGAGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGATGAGTCTCCA 783  
Db 725 CTCGCCAGAGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGATGAGTCTCCA 784  
Qy 784 CTCGCCAGAGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGATGAGTCTCCA 843  
Db 785 CTCGCCAGAGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGATGAGTCTCCA 844  
Qy 844 CTCGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGGCTCGATCA 903  
Db 845 CTCGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGGCTCGATCA 904  
Qy 904 GAACTCCAAACTGATCAGCCCCCTTGCAATGTCTGAAACCAAGTGTGAAACTGCACCAAC 963  
Db 905 GAACTCCAAACTGATCAGCCCCCTTGCAATGTCTGAAACCAAGTGTGAAACTGCACCAAC 964  
Qy 964 CCCGAGGAGGGCCCTGCGCCCTGCCACGCAACCCCTTCCCTATAGCAGATGCTCT 1023  
Db 965 CCCGAGGAGGGCCCTGCGCCCTGCCACGCAACCCCTTCCCTATAGCAGATGCTCT 1024  
Qy 1024 CATCCCTTCCCATTCACCCCTCTCCAGCCCTGGAAACCTCACCCCTCTGGAGTCCCTTCCTG 1083  
Db 1025 CATCCCTTCCCATTCACCCCTCTCCAGCCCTGGAAACCTCACCCCTCTGGAGTCCCTTCCTG 1084  
Qy 1084 GGCAACTGGCCTGTGTAGACAGCAGAAACCCCTTGCCCTGACCCCACTGTAGCAAACTG 1143  
Db 1085 GGCAACTGGCCTGTGTAGACAGCAGAAACCCCTTGCCCTGACCCCACTGTAGCAAACTG 1144  
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGGAGCCAGCTGCTGTCT 1203  
Db 1145 GCCTGTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGGAGCCAGCTGCTGTCT 1204  
Qy 1204 CGTGTGCCCCTAGAGAAATTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1263  
Db 1205 CGTGTGCCCCTAGAGAAATTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1264  
Qy 1264 GTGAGTCAAGCCAGGCCACAGCTGACAGCTTGGCCAAAGACCTTGGGACGACGCGGC 1323  
Db 1265 GTGAGTCAAGCCAGGCCACAGCTGACAGCTTGGCCAAAGACCTTGGGACGACGCGGC 1324  
Qy 1324 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACAGAGGCTGCTGCTCACTGAG 1383  
Db 1325 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACAGAGGCTGCTGCTCACTGAG 1384  
Qy 1384 AAACCTCAAGCCAGTGGATATGAGTACCGAGAGAAAGTCCACTGGGCCACGACCAAGCTC 1443  
Db 1385 AAACCTCAAGCCAGTGGATATGAGTACCGAGAGAAAGTCCACTGGGCCACGACCAAGCTC 1444  
Qy 1444 CGCCTGGGAGAGGCTCTCTTCGGAGAGGTGCAAGATGGAGGACAAAGCAGACTGGCTTC 1503  
Db 1445 CGCCTGGGAGAGGCTCTCTTCGGAGAGGTGCAAGATGGAGGACAAAGCAGACTGGCTTC 1504  
Qy 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTCCGGCAGAGAGCTGATGCAATGT 1563  
Db 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTCCGGCAGAGAGCTGATGCAATGT 1564  
Qy 1564 GCAGGATTCACCTCAACCAAGATTTGCTTGTATGAGCTGTGTGAGAGAGGGCTTGG 1623  
Db 1565 GCAGGATTCACCTCAACCAAGATTTGCTTGTATGAGCTGTGTGAGAGAGGGCTTGG 1624  
Qy 1624 GTCAACATCTTCATGAGCTGTGAAAGGTGCTCTTCGGGCCAGCTGTGTCAAGAGCAG 1683  
Db 1625 GTCAACATCTTCATGAGCTGTGAAAGGTGCTCTTCGGGCCAGCTGTGTCAAGAGCAG 1684

Qy 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTGTACTCTGGGCCAGGCCCTGGAGGCTCTGGAA 1743  
Db 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTGTACTCTGGGCCAGGCCCTGGAGGCTCTGGAA 1744  
Qy 1744 TACCTCCACTCACGAAGGATTTCTGCATGGGGAACGTCAAAAGCTGACAAAGTGTCTCTGTCC 1803  
Db 1745 TACCTCCACTCACGAAGGATTTCTGCATGGGGAACGTCAAAAGCTGACAAAGTGTCTCTGTCC 1804  
Qy 1804 AGCGATGGAGGCCACGACGCCCTCTGTGACTTTTGGCCATGTGTGTGTCTTTCAACCTGAT 1863  
Db 1805 AGCGATGGAGGCCACGACGCCCTCTGTGACTTTTGGCCATGTGTGTGTCTTTCAACCTGAT 1864  
Qy 1864 GSCCTGGGAAAGTCCCTTCTCACAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCT 1923  
Db 1865 GSCCTGGGAAAGTCCCTTCTCACAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCT 1924  
Qy 1924 CCGGAGGTGTGTCTGGGACAGGAGCTGCGACGCCCAAGGTGTGTCTGGAGCAGCTGTCT 1983  
Db 1925 CCGGAGGTGTGTCTGGGACAGGAGCTGCGACGCCCAAGGTGTGTCTGGAGCAGCTGTCT 1984  
Qy 1984 ATGATGTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC 2043  
Db 1985 ATGATGTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC 2044  
Qy 2044 TGCTTCAAGATTTGCCAGCGAGCTCCGCTGTGTAGGGAGATCCCAACCTCTCTCGCCCT 2103  
Db 2045 TGCTTCAAGATTTGCCAGCGAGCTCCGCTGTGTAGGGAGATCCCAACCTCTCTCGCCCT 2104  
Qy 2104 CTCACAGCCCAGGCCATCCAGAGGGGTGAGGAAAGGCCCATCCACCGGTGTCTGCA 2163  
Db 2105 CTCACAGCCCAGGCCATCCAGAGGGGTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164  
Qy 2164 GCGAGCTGGAGAGGAAGGTGAAACGGGCACTACAGCAAGTGGAGGTCTGAGAGCCCT 2223  
Db 2165 GCGAGCTGGAGAGGAAGGTGAAACGGGCACTACAGCAAGTGGAGGTCTGAGAGCCCT 2224  
Qy 2224 TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTAACCAACAG 2283  
Db 2225 TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCAAAATCAAGCCAAATTAACCAACAG 2284  
Qy 2284 ACCCTTCCATGCCAGCCGAGAGAGCTTTCCGCCAAGGGCCCCCAGGGCCCCCGCCAGCTGAG 2343  
Db 2285 ACCCTTCCATGCCAGCCGAGAGAGCTTTCCGCCAAGGGCCCCCAGGGCCCCCGCCAGCTGAG 2344  
Qy 2344 GAGACAAAGCCAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAGCCCCCAGAGCCA 2403  
Db 2345 GAGACAAAGCCAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAGCCCCCAGAGCCA 2404  
Qy 2404 AACAGTCTCTCTCCCTTCACTTTTGAGCAAGAGGAGTCTGGGATGTGGGAACCTTACCT 2463  
Db 2405 AACAGTCTCTCTCCCTTCACTTTTGAGCAAGAGGAGTCTGGGATGTGGGAACCTTACCT 2464  
Qy 2464 CTGTCTCTCTGGAGCCAGCCCTGCCAGAAAACCCAGCTCACAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCTCTCTGGAGCCAGCCCTGCCAGAAAACCCAGCTCACAGAGCGGAAAGCAACC 2524  
Qy 2524 GTCCCGGAGCAGGAACCTCAGCAGCTGGAATATGAAATTTATTCCTCAAAGCTGTCTCCAG 2583  
Db 2525 GTCCCGGAGCAGGAACCTCAGCAGCTGGAATATGAAATTTATTCCTCAAAGCTGTCTCCAG 2584  
Qy 2584 CCATTTTCTCTGGAGGACGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCTCTCC 2643  
Db 2585 CCATTTTCTCTGGAGGACGAGCAAAATTTCTCTGCTGCTCAGCATCGACAGCTCTCC 2644  
Qy 2644 CTGTCCGATGACAGTGAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCTCTG 2703  
Db 2645 CTGTCCGATGACAGTGAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCTCTG 2704  
Qy 2704 AGCTCAGGGGTACTCTCTGGAGCAGCCAGGGCTCGAAGCTCCAGCTGGAGACATG 2763  
Db 2705 AGCTCAGGGGTACTCTCTGGAGCAGCCAGGGCTCGAAGCTCCAGCTGGAGACATG 2764

QY	2764	GTGCTGGCCCGGGGGGGCCCAACGACACACCCCAAGCTATTTCAATGGTGTGAAGTCCAA	2823
Db	2765	GTGCTGGCCCGGGGGGGCCCAACGACACACCCCAAGCTATTTCAATGGTGTGAAGTCCAA	2824
QY	2824	ATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA	2883
Db	2825	ATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA	2884
QY	2884	GACATGGCCACTGGGATCAGCAGCCAGATCCAGCTGCAGCCTTTCAGCTTGGTCAACCAA	2943
Db	2885	GACATGGCCACTGGGATCAGCAGCCAGATCCAGCTGCAGCCTTTCAGCTTGGTCAACCAA	2944
QY	2944	GACGGGACGCTGTTGCTACGACATGAGAGTGCAGACTGGGATCGACCTGCAGTGC	3003
Db	2945	GACGGGACGCTGTTGCTACGACATGAGAGTGCAGACTGGGATCGACCTGCAGTGC	3004
QY	3004	ACACTGGCCCTTGATGGGAGCTTCCGCTGGAGCTCGAGGGTCAAGCATGGCCAGCTGGAG	3063
Db	3005	ACACTGGCCCTTGATGGGAGCTTCCGCTGGAGCTCGAGGGTCAAGCATGGCCAGCTGGAG	3064
QY	3064	AACAGGCCCTAAACCTGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCTTTCCTG	3123
Db	3065	AACAGGCCCTTAACCTGCTCCACCGCGGCTCCACACTGCGGAAAGCAGCCTTTCCTG	3124
QY	3124	CTCGTGCAGATGCTGCTGCTGAAACACAGGCTCAGCCGTTCCAGGGGATTCAGGCC	3183
Db	3125	CTCGTGCAGATGCTGCTGCTGAAACACAGGCTCAGCCGTTCCAGGGGATTCAGGCC	3184
QY	3184	CCCCGGCTCACAGTGGGAAACAGGGCTCGAGCAGCAGGTTGGGGCAAGCAGATGCC	3243
Db	3185	CCCCGGCTCACAGTGGGAAACAGGGCTCGAGCAGCAGGTTGGGGCAAGCAGATGCC	3244
QY	3244	TCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGAAACCAATCCCGCAGTGAA	3303
Db	3245	TCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGAAACCAATCCCGCAGTGAA	3304
QY	3304	GAGACAGAGGAGTGGCAGAGTTACTGGGGAACAAACAGGGATCTTTTCTGCC	3363
Db	3305	GAGACAGAGGAGTGGCAGAGTTACTGGGGAACAAACAGGGATCTTTTCTGCC	3364
QY	3364	CCTGCTCAGTTCAGTTCGCTGACCCGCTTGATCAGTGACCAATTTGTTGGCAGACAGG	3423
Db	3365	CCTGCTCAGTTCAGTTCGCTGACCCGCTTGATCAGTGACCAATTTGTTGGCAGACAGG	3424
QY	3424	GGAGAGCAGCTTCAGACCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTTCACCCCTC	3483
Db	3425	GGAGAGCAGCTTCAGACCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTTCACCCCTC	3484
QY	3484	CAGGCTGCTGTAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTCAAGTGCAGAACCA	3543
Db	3485	CAGGCTGCTGTAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTCAAGTGCAGAACCA	3544
QY	3544	GGTCAGCAGGTATGCCCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTCGCTGGCCCTC	3603
Db	3545	GGTCAGCAGGTATGCCCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTCGCTGGCCCTC	3604
QY	3604	ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCTGACCCCTGGGAAGCT	3663
Db	3605	ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCTGACCCCTGGGAAGCT	3664
QY	3664	CCCTGGTGAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTCACCTCT	3723
Db	3665	CCCTGGTGAATACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTCACCTCT	3724
QY	3724	GGTTCAGCCAAAGGACCAAGAGTGTGTAAGTGAAGTGTCTCAGTCCCGCAGCATGT	3783
Db	3725	GGTTCAGCCAAAGGACCAAGAGTGTGTAAGTGAAGTGTCTCAGTCCCGCAGCATGT	3784
QY	3784	GCCCTTGTGCTGGCTACCACTCTTCCAGAGCAGCAGGCCCGCAGGCCCTTTCAGGC	3843
Db	3785	GCCCTTGTGCTGGCTACCACTCTTCCAGAGCAGCAGGCCCGCAGGCCCTTTCAGGC	3844
QY	3844	CCAGCACTGCCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGAT	3903

Db	3845	CCAGCACTGCCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGAT	3904
QY	3904	GCAGGATATGCTGCACAGAAACAGTCTGTGATGGACATGATCAGTGTAAAGGAAAGCAG	3963
Db	3905	GCAGGATATGCTGCACAGAAACAGTCTGTGATGGACATGATCAGTGTAAAGGAAAGCAG	3964
QY	3964	CAGAGAGAGCGTCCGGCGCCCCAGCCCACTATCAGTGTCCAGCGTCTGGTTCCCGCAG	4023
Db	3965	CAGAGAGAGCGTCCGGCGCCCCAGCCCACTATCAGTGTCCAGCGTCTGGTTCCCGCAG	4024
QY	4024	AGCAGAGCTCAGCATCAGACTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGC	4083
Db	4025	AGCAGAGCTCAGCATCAGACTGACACTCACCTGCCCTGCCCTGGCCAGAGGGTACTGC	4084
QY	4084	CGAGCGCACTTTGACATCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTCTGGCAGGGC	4143
Db	4085	CGAGCGCACTTTGACATCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTCTGGCAGGGC	4144
QY	4144	AGGCGAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4203
Db	4145	AGGCGAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4204
QY	4204	GTCTTGAGCTGTCCACATGATGACTCTCTCAAAACCTTCTCCAGATTTCTTAAGATA	4263
Db	4205	GTCTTGAGCTGTCCACATGATGACTCTCTCAAAACCTTCTCCAGATTTCTTAAGATA	4264
QY	4264	GCACCCCTTCCCATTTGCCCTTAGCCCTTCTCCAGGGAGCTACTCAGGACTC	4323
Db	4265	GCACCCCTTCCCATTTGCCCTTAGCCCTTCTCCAGGGAGCTACTCAGGACTC	4324
QY	4324	ACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCCTCAACCTCTGGG	4383
Db	4325	ACGTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCCTCAACCTCTGGG	4384
QY	4384	GCAGGGACGCCGAGACTCGTGGGAGAGCTCATTTCCACATCTTGGCCAGACAGCCTT	4443
Db	4385	GCAGGGACGCCGAGACTCGTGGGAGAGCTCATTTCCACATCTTGGCCAGACAGCCTT	4444
QY	4444	TGTCAGCTGTCCACATTTGAGTTCAGCTGTCTCCCGGGAGAGAGCCCGGCCCGCCAGCAC	4503
Db	4445	TGTCAGCTGTCCACATTTGAGTTCAGCTGTCTCCCGGGAGAGAGCCCGGCCCGCCAGCAC	4504
QY	4504	ATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA	4563
Db	4505	ATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA	4564
QY	4564	TAAAGTTTGGGGTGAATGACAAATGTTAAAAA	4595
Db	4565	TAAAGTTTGGGGTGAATGACAAATGTTAAAAA	4596

RESULT 5  
ADR25114  
ID ADR25114 standard; DNA; 4596 BP.  
XX ADR25114;  
AC ADR25114;  
XX 21-OCT-2004 (first entry)  
XX  
XX Breast cancer prognosis marker #975.  
DE ds; breast cancer; prognosis; gene expression; diagnosis.  
XX  
XX Homo sapiens.  
XX  
XX WO2004065545-A2.  
XX  
XX 05-AUG-2004.  
XX  
XX 15-JAN-2004; 2004WO-US001100.  
PF  
XX 15-JAN-2003; 2003US-00342887.  
PR

XX	(ROSE-) ROSETTA INPHARMATICS LLC.
PA	(NECA-) NETHERLANDS CANCER INST.
XX	
XX	Van't Veer LJ, He Y;
XX	WPI; 2004-593473/57.
DR	
XX	Classifying a breast cancer patient according to prognosis comprises
PT	determining the similarity between the level of expression of each of
PT	five genes in a cell sample taken from patient, to control levels.
XX	
XX	Disclosure; SEQ ID NO 975; 226pp; English.
XX	
CC	The invention relates to a method of classifying a breast cancer patient
CC	according to prognosis by determining the similarity between the level of
CC	expression of each of five genes for which markers are listed in the
CC	specification, in a cell sample taken from the breast cancer patient, to
CC	control levels of expression for each respective five genes to obtain a
CC	patient similarity value. The methods are useful for classifying a breast
CC	cancer patient according to prognosis. Kits and computer program products
CC	are useful for data analysis using the diagnostic, prognostic and
CC	statistical methods of the invention. This sequence corresponds to a
CC	marker used in the method of the invention.
XX	
SQ	Sequence 4596 BP; 1055 A; 1401 C; 1299 G; 841 T; 0 U; 0 Other;
Query Match 99.6%; Score 4578.4; DB 13; Length 4596;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
QY	5 GGGGACCTGTCGGTGTGGAAAGTGTAGCTGTGA-AGGTGGACTCTGTACCAATTGAGG 63
DB	5 GGGGACCTGTCGGTGTGGAAAGTGTAGCTGTGAGAGGTGGACTCTGTACCAATTGAGG 64
QY	64 ATGTTTGGAGATAGTATGTGTGCAGAGGCACACATAACAGCAGACAGCCCTTTGCC 123
DB	65 ATGTTTGGAGATAGTATGTGTGCAGAGGCACACATAACAGCAGACAGCCCTTTGCC 124
QY	124 CCTGCCTTCTCCCAACCAAGGCTGACCTGTCTCTCCAGGTCTGGGATTTCTAAGT 183
DB	125 CCTGCCTTCTCCCAACCAAGGCTGACCTGTCTCTCCAGGTCTGGGATTTCTAAGT 184
QY	184 GACCTGCTCTGTGTTGGTCTCTCTCAGATGAGCACAAAGCTGGAGATGCGATGATG 243
DB	185 GACCTGCTCTGTGTTGGTCTCTCTCAGATGAGCACAAAGCTGGAGATGCGATGATG 244
QY	244 GAAATGGCTGCCAGGTGCCCTGGCTCAGAGTGGGCGACAGAAGAACTCCCCAAG 303
DB	245 GAAATGGCTGCCAGGTGCCCTGGCTCAGAGTGGGCGACAGAAGAACTCCCCAAG 304
QY	304 CCAAGAGGAAGACGCCCGCACTGGGGAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 363
DB	305 CCAAGAGGAAGACGCCCGCACTGGGGAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 364
QY	364 GTGGAGAGAGCCCTGTGTTCTTCGGAAGTGGAGATCCTGAATGACGTGATTACCAG 423
DB	365 GTGGAGAGAGCCCTGTGTTCTTCGGAAGTGGAGATCCTGAATGACGTGATTACCAG 424
QY	424 GGCACAGCCCAAGGAGGCTCCGAGGCAGGCCTGCCATCTCTATCATCGCCCAAGGCT 483
DB	425 GGCACAGCCCAAGGAGGCTCCGAGGCAGGCCTGCCATCTCTATCATCGCCCAAGGCT 484
QY	484 GAGTGTGAAATAGCCAAAGATTTCAGCCCCACCTTTTCAGAACGCAATTTCACTGCTGG 543
DB	485 GAGTGTGAAATAGCCAAAGATTTCAGCCCCACCTTTTCAGAACGCAATTTCACTGCTGG 544
QY	544 TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCAATGCT 603
DB	545 TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCAATGCT 604
QY	604 ACAGAGGGCAAAATGGCCCGTGTGTGTGGAGGGGAAAGCGTCGACAGCAAGCCCGAAG 663





Db 3905 GCAGGATATGCTGACAGAAAGTCTGTGGATGACATGATCACTGTAAAGAAAGCAG 3964  
Qy 3964 CAGAGAGAGCTCGCGGCCCCAGCCCACTATCAGTGTCCAGCGTCTGTTCCCGCAG 4023  
Db 3965 CAGAGAGAGCTCGCGGCCCCAGCCCACTATCAGTGTCCAGCGTCTGTTCCCGCAG 4024  
Qy 4024 AGCAGAGCTGACATCAGACTGACACTCACCTGCTGCCCTGCCCTGCCAGAGGCTACTGC 4083  
Db 4025 AGCAGAGCTGACATCAGACTGACACTCACCTGCTGCCCTGCCCTGCCAGAGGCTACTGC 4084  
Qy 4084 CGACGGCACTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTCTGGCAGGCG 4143  
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCCCTCTGGCAGGCG 4144  
Qy 4144 AGGCGAGGCGAGTACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4203  
Db 4145 AGGCGAGGCGAGTACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4204  
Qy 4204 GTCTTGAGCTGCCACATGATGATGCTCCTCAAACTCTTCCAGATTTCTTAAGAATA 4263  
Db 4205 GTCTTGAGCTGCCACATGATGATGCTCCTCAAACTCTTCCAGATTTCTTAAGAATA 4264  
Qy 4264 GCACCCCTTCCCACTGCCCCAGCTTAGCCCTCTTCTCCAGGGAGCTACTCAGGACTC 4323  
Db 4265 GCACCCCTTCCCACTGCCCCAGCTTAGCCCTCTTCTCCAGGGAGCTACTCAGGACTC 4324  
Qy 4324 AGCTAGCAATTAATCAGCTGTGAATCGTCAGGGGTGTCTGATGCTCAACCTCCTGGG 4383  
Db 4325 AGCTAGCAATTAATCAGCTGTGAATCGTCAGGGGTGTCTGATGCTCAACCTCCTGGG 4384  
Qy 4384 CAGGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTCCAGAGAGCCTT 4443  
Db 4385 CAGGGGAGCGGAGACTCCGTGGGAGAGCTCATTTCCACATCTTCCAGAGAGCCTT 4444  
Qy 4444 TGTCAGCTGTCCACATGAGTCACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAC 4503  
Db 4445 TGTCAGCTGTCCACATGAGTCACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAC 4504  
Qy 4504 ATAAAGAACTGAGCTGGTGTACTGAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAA 4563  
Db 4505 ATAAAGAACTGAGCTGGTGTACTGAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAA 4564  
Qy 4564 TAAAGTTGGGTGATGACAAATGTTAAAAA 4595  
Db 4565 TAAAGTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 6  
ACN44439  
ID ACN44439 standard; cDNA; 4684 BP.  
XX  
AC ACN44439;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human mRNA sequence hCT1952652.  
XX  
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO2003073826-A2.  
XX  
PD 12-SEP-2003.  
XX  
PF 28-FEB-2003; 2003WO-US006235.  
XX  
PR 01-MAR-2002; 2002US-00087192.  
XX  
PA (SAGR-) SAGRES DISCOVERY.  
XX  
PI Morris DW;

XX WPI; 2003-328604/31.  
XX  
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
XX  
PS Claim 1; SEQ ID NO 887; Opp; English.  
XX  
CC The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for  
CC screening drug candidates; (ii) for screening of bioactive agent capable  
CC of binding to Carcinoma Associated protein (CAP); (iii) for screening of  
CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
CC carcinoma including lymphoma. The present sequence is one such CA gene  
CC sequence. Note: This patent is an equivalent to basic patent  
CC US2002182586A1, for which no sequence data was published  
XX  
SQ Sequence 4684 BP; 1057 A; 1434 C; 1328 G; 865 T; 0 U; 0 Other;  
Query Match 96.6%; Score 4440; DB 11; Length 4684;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 4549; Conservative 0; Mismatches 30; Indels 11; Gaps 8;  
Qy 3 CGGGGGGACTGTCCGTGTGGAAGCTGTAGCTGTGA-AGGTGGAAGCTGTGACATTGA 61  
Db 53 CGGGGGGACTGTCCGTGTGGAAGCTGTAGCTGTGAGAGGTGGAAGCTGTGACATTGA 112  
Qy 62 GGATGTTGGAGGATGATATGTGGCAGAGGACACATAAAGGAGAGACCTTTG 121  
Db 113 GGATGTTGGAGGATGATATGTGGCAGAGGACACATAAAGGAGAGACCTTTG 172  
Qy 122 CCCCTGCTTCTCCCAAGCTGACCTGTGTCTTCCAGGTCTGGGATTTCTAA 181  
Db 173 CCCCTGCTTCTCCCAAGCTGACCTGTGTCTTCCAGGTCTGGGATTTCTAA 232  
Qy 182 GTGACCTGCTGTGTGTTGGTCTCTCTCAGGATGAGCACAAGCTGGAGATGGCAGTGA 241  
Db 233 GTGACCTGCTGTGTGTTGGTCTCTCTCAGGATGAGCACAAGCTGGAGATGGCAGTGA 292  
Qy 242 TGGAAATGGCTTCCCGAGGTGCTGCTCAGAGTGGGGCAGAGAGGAACTCCCA 301  
Db 293 TGGAAATGGCTTCCCGAGGTGCTGCTCAGAGTGGGGCAGAGAGGAACTCCCA 352  
Qy 302 AGCCAAAGGAGAGAGCGCCGCTGCGGGAAGAAACAGAGCTCCGCTTACAAGCTTTGAGG 361  
Db 353 AAGCCAAAGGAGAGAGCGCCGCTGCGGGAAGAAACAGAGCTCCGCTTACAAGCTTTGAGG 412  
Qy 362 CCGTGGAGAGAGCGCTGTGTTCGGAAGTGGGAGATCTGTAATGACGTGATACCA 421  
Db 413 CCGTGGAGAGAGCGCTGTGTTCGGAAGTGGGAGATCTGTAATGACGTGATACCA 472  
Qy 422 AGGCGACAGCCAAAGGAGCTCCGAGGAGGCGGAGCTGCCATCTCTATCATCGCCAGG 481  
Db 473 AGGCGACAGCCAAAGGAGCTCCGAGGAGGCGGAGCTGCCATCTCTATCATCGCCAGG 532  
Qy 482 CTGAGTGTGAGAAATAGCAAGAGTTTCCAGCCCACTTTTCAGAAAGCTTTTCATCGCTG 541  
Db 533 CTGAGTGTGAGAAATAGCAAGAGTTTCCAGCCCACTTTTCAGAAAGCTTTTCATCGCTG 592  
Qy 542 GGTCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATG 601  
Db 593 GGTCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATG 652  
Qy 602 CTACAGAGGCAAAATGGCCCGTGTGTGGAGGGAAGCGTCGACGCAAGCCCGA 661  
Db 653 CTACAGAGGCAAAATGGCCCGTGTGTGGAGGGAAGCGTCGACGCAAGCCCGA 712







PI Wu L, Rothe M;  
XX WPI: 1999-094902/08.  
XX  
PT Screening agents for modulating interaction of nuclear factor kappaB  
PT inducing kinase - with kinase-binding target, useful for controlling  
PT levels of the kinase, for treatment and diagnosis of conditions  
PT associated with e.g. inhibition of signal transduction by tumour necrosis  
XX factor.  
XX  
PS Disclosure; Col 11-14; 16pp; English.  
XX  
CC This sequence encodes a novel human nuclear factor-kappaB (NFkB) -inducing  
CC kinase (NIK) which is used in a method for screening for agents that  
CC modulate the interaction of NIK with a NIK-binding target. The encoded  
CC protein can be used as a modulator of cellular functions at the NIK  
CC level, or for development of such compounds. NFkB is involved in  
CC expression of many immune and inflammatory responses and of some  
CC important viral genes. The protein may be used diagnostically and  
CC therapeutically, in conditions associated with abnormal utilisation of  
CC pathways that involve NFkB, e.g. inhibition of signal transduction by  
CC tumour necrosis factor (TNF)  
XX  
SQ Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 68.5%; Score 3149.6; DB 2; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 232 ATGGCAGTGATGGAATGCGTCCAGGTGCGCCCTGAGTCCCTGAGTGGGGCAGCAGAAAG 291  
DB 1 ATGGCAGTGATGGAATGCGTCCAGGTGCGCCCTGAGTGGGGCAGCAGAAAG 60

QY 292 GAACTCCCAAGCAGGAG 351  
DB 61 GAACTCCCAAGCAGGAG 120

QY 352 AAGCTTGAGGCGGTGGAG 411  
DB 121 AAGCTTGAGGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

QY 412 GTGATTACCAAGGGCAG 471  
DB 181 GTGATTACCAAGGGCAG 240

QY 472 ATCGCCAGGCTGAGTGAG 531  
DB 241 ATCGCCAGGCTGAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 532 TTGATCGTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAAT 591  
DB 301 TTGATCGTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAA 360

QY 592 GTGGCCCATGTACAGAGGGCAAAATGCGCGTGTGTGGAGAGGAGAGAGAGAGAGAGAG 651  
DB 361 GTGGCCCATGTACAGAGGGCAAAATGCGCGTGTGTGGAGAGGAGAGAGAGAGAGAG 420

QY 652 AAAGCCGGAAG 711  
DB 421 AAAGCCGGAAG 480

QY 712 TTGGCCAAACCCCTCCCAAG 771  
DB 481 TTGGCCAAACCCCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 772 GATGAGTCTCACTCGGCGGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCCTG 831  
DB 541 GATGAGTCTCACTCGGCGGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCCTG 600

QY 832 AAGGAAACAGGCGCTTGGCACTCTGTTTAAAGCAGCTTGGCGAGGGGCTTACGCGCGCT 891  
DB 601 AAGGAAACAGGCGCTTGGCACTCTGTTTAAAGCAGCTTGGCGAGGGGCTTACGCGCGCT 660

QY 892 CTGCTCGATCAGAACTCCCAAACTGATCAGCCCTTTGCAATGTGTGAACCAACGCTGTGG 951  
DB 661 CTGCTCGATCAGAACTCCCAAACTGATCAGCCCTTTGCAATGTGTGAACCAACGCTGTGG 720

QY 952 AAATGTCACCAACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1011  
DB 721 AAATGTCACCAACCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780

QY 1012 AGCAGAGTGCCTCATCCCTTTCCCATTCACCTCTCCAGCCCTCGAAACCTTCACCTCTG 1071  
DB 781 AGCAGAGTGCCTCATCCCTTTCCCATTCACCTCTCCAGCCCTCGAAACCTTCACCTCTG 840

QY 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTGGCTGTGACCCAC 1131  
DB 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTGGCTGTGACCCAC 900

QY 1132 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTCGAGGCC 1191  
DB 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTCGAGGCC 960

QY 1192 AGCTGCTGTCTCGTGTGCGCCATGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1251  
DB 961 AGCTGCTGTCTCGTGTGCGCCATGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1020

QY 1252 CTGCAAGGCGGTGAGCTCAAGCCAGGCCACAGCTGACCCAGCTGAGCCGCAAGACCTGG 1311  
DB 1021 CTGCAAGGCGGTGAGCTCAAGCCAGGCCACAGCTGACCCAGCTGAGCCGCAAGACCTGG 1080

QY 1312 GCAGCAGGGGCTCCAGATCCCGGAGGCCAGCCCAAACTGAGGACACAGAGGGTGTG 1371  
DB 1081 GCAGCAGGGGCTCCAGATCCCGGAGGCCAGCCCAAACTGAGGACACAGAGGGTGTG 1140

QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTTGGGCC 1431  
DB 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTTGGGCC 1200

QY 1432 ACGCACAGCTCGGCTGGGCGAGAGGCTCTTCCGAGAGGTGTCACAGATGGAGGACAAG 1491  
DB 1201 ACGCACAGCTCGGCTGGGCGAGAGGCTCTTCCGAGAGGTGTCACAGATGGAGGACAAG 1260

QY 1492 CAGACTGGCTTCAGTCCGCTGTCAAAAGGTCGCGCTGGAAGTATTTTCGGGCGAGAGAG 1551  
DB 1261 CAGACTGGCTTCAGTCCGCTGTCAAAAGGTCGCGCTGGAAGTATTTTCGGGCGAGAGAG 1320

QY 1552 CTGATGTCATGTGAGGATTTGACCTCAACCCAGAAATGTCCTTTGTATGAGAGTGTGAGA 1611  
DB 1321 CTGATGTCATGTGAGGATTTGACCTCAACCCAGAAATGTCCTTTGTATGAGAGTGTGAGA 1380

QY 1612 GAAGGGCTTTGGGTCAAACATCTTTATGAGAGTGTGGAAGGTGGCTCCCTGGGCGCAGCTG 1671  
DB 1381 GAAGGGCTTTGGGTCAAACATCTTTATGAGAGTGTGGAAGGTGGCTCCCTGGGCGCAGCTG 1440

QY 1672 GTCAAGAGAGAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACCTTGGGCGCAGGCCCTG 1731  
DB 1441 GTCAAGAGAGAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACCTTGGGCGCAGGCCCTG 1500

QY 1732 GAGGCTCTGGAATACCTCCACTCAAGAGGATTTCTGATGGGAGCTCAAAAGCTGACAAAC 1791  
DB 1501 GAGGCTCTGGAATACCTCCACTCAAGAGGATTTCTGATGGGAGCTCAAAAGCTGACAAAC 1560

QY 1792 GTGCTCTGTCCAGCGATGGGAGCCACGAGCCCTCTGTGATCTTTGGCCATGTGTGTGT 1851  
DB 1561 GTGCTCTGTCCAGCGATGGGAGCCACGAGCCCTCTGTGATCTTTGGCCATGTGTGTGT 1620

QY 1852 CTTCAACCTGATGGCTGGGAAAGTCTTCTGCTCAAGGAGACTACATCCCTGGGCGACAGAG 1911  
DB 1621 CTTCAACCTGATGGCTGGGAAAGTCTTCTGCTCAAGGAGACTACATCCCTGGGCGACAGAG 1680

QY 1912 ACCCACATGCTCGGAGGTGTGCTGGGCGAGGAGCTGCGAGCCCAAGAGTGGATGTCTGG 1971  
DB 1681 ACCCACATGCTCGGAGGTGTGCTGGGCGAGGAGCTGCGAGCCCAAGAGTGGATGTCTGG 1740

QY 1972 AGCAGTGTGTATGATGTGCAATGCTCAAAGGCTGCCAGCCCTGGACTCAGTTCTTTC 2031

1741 AGCAGCTGCTGATGATGCTGCAATGCTCAAGGCTGCCACCCCTGACTCAGTTCTTC 1800  
2032 CGAGGCGCTGCTGCTCAAGATTGCCAGGAGCCCTCCGCTGTGAGGAGATCCACCC 2091  
1801 CGAGGCGCTGCTGCTCAAGATTGCCAGGAGCCCTCCGCTGTGAGGAGATCCACCC 1860  
2092 TCCTGGGCGCTGCTCAGAGCCAGGCGCATCCAGAGGGGCTGAGAAAGAGCCCATCCAC 2151  
1861 TCCTGGGCGCTGCTCAGAGCCAGGCGCATCCAGAGGGGCTGAGAAAGAGCCCATCCAC 1920  
2152 CGCTGTCTGACCGGAGCTGGGAGGAGTGAACCGGCACTACAGCAAGTGGAGGT 2211  
1921 CGCTGTCTGACCGGAGCTGGGAGGAGTGAACCGGCACTACAGCAAGTGGAGGT 1980  
2212 CTGAAGAGCCCTTGAGGGGAGAAATAAAGAAACAAGACATCCACCCCAATCAAGCC 2271  
1981 CTGAAGAGCCCTTGAGGGGAGAAATAAAGAAACAAGACATCCACCCCAATCAAGCC 2040  
2272 AATTACCACAGACCTTCNTGCCAGCGGAGAGCTTCGCCAAGGGCCCGCAGGGCC 2331  
2041 AATTACCACAGACCTTCNTGCCAGCGGAGAGCTTCGCCAAGGGCCCGCAGGGCC 2100  
2332 CGGCCAGCTGAGGAGCAACAGCGAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAG 2391  
2101 CGGCCAGCTGAGGAGCAACAGCGAGAGCCCTTAAGCTCCAGCCTCCTCTCCACAGAG 2160  
2392 CCCCCAGAGCCAAACAAGTCTCTCTCTTGACTTTGAGCAAGGAGAGTCTGGAGTGTGG 2451  
2161 CCCCCAGAGCCAAACAAGTCTCTCTCTTGACTTTGAGCAAGGAGAGTCTGGAGTGTGG 2220  
2452 GAACCTTACCTGTCTCTCTGAGCAGCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2511  
2221 GAACCTTACCTGTCTCTCTGAGCAGCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2280  
2512 CGGAAGCAACCGTCCCGAGCAGGAACTGCGAGAGTGGAAATAGAATATTCTCTCAAC 2571  
2281 CGGAAGCAACCGTCCCGAGCAGGAACTGCGAGAGTGGAAATAGAATATTCTCTCAAC 2340  
2572 AGCTGTCCAGCATTCTCTGAGAGAGAGGAGAAATTTCTCTGTGCTCAGCATC 2631  
2341 AGCTGTCCAGCATTCTCTGAGAGAGAGGAGAAATTTCTCTGTGCTCAGCATC 2400  
2632 GACAGCTCTCCCTGCTGAGTACAGTGAAGAGACCCATCAAGGCTCTCAAGAGCTCG 2691  
2401 GACAGCTCTCCCTGCTGAGTACAGTGAAGAGACCCATCAAGGCTCTCAAGAGCTCG 2460  
2692 CGGGACACCTGAGCTCAGGCGGTACACTCTCTGAGCAGCCAGGCGGAGCTCGAAGCTCC 2751  
2461 CGGGACACCTGAGCTCAGGCGGTACACTCTCTGAGCAGCCAGGCGGAGCTCGAAGCTCC 2520  
2752 AGCTGGAACATGCTGCTGGCGGGGGCGCCACCGACACCCCAAGCTATTTCATGCT 2811  
2521 AGCTGGAACATGCTGCTGGCGGGGGCGCCACCGACACCCCAAGCTATTTCATGCT 2580  
2812 GTCAAGTCCAAATACAGTCTCTTAATGTGTAACCTTCACATCGGGAGTTTCCACCGG 2871  
2581 GTCAAGTCCAAATACAGTCTCTTAATGTGTAACCTTCACATCGGGAGTTTCCACCGG 2640  
2872 GTCAAGTGGGAGACATCGCACTGCGATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2931  
2641 GTCAAGTGGGAGACATCGCACTGCGATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2700  
2932 TTGGTCCAAAGACGGGAGCCCTGTCGTACGATGAGGTGCGAGCTCGGGCATC 2991  
2701 TTGGTCCAAAGACGGGAGCCCTGTCGTACGATGAGGTGCGAGCTCGGGCATC 2760  
2992 GACTCGAGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 3051  
2761 GACTCGAGTGACACTGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 2820  
3052 GGCCAGCTGGAGAACAGGCCCTTAACCTTCACCGGCTCCACCGGCTCCACACTCGCGAAA 3111

2821 GGCCAGCTGGAGAACAGGCCCTTAACCCCTGCCCTCCACCGCGGCTCCACACTGCGCGAAA 2880  
3112 GCAGCCTTCTGCTCGGTGCAATGCTGCGCTGAAACACAGGCTCAGCCGTTCCAGG 3171  
2881 GCAGCCTTCTGCTCGGTGCAATGCTGCGCTGAAACACAGGCTCAGCCGTTCCAGG 2940  
3172 GGATTGCGAGCCCCCGGCTCACAGTGGGAACACAGGCTCGCAGCAAGGTGGGGC 3231  
2941 GGATTGCGAGCCCCCGGCTCACAGTGGGAACACAGGCTCGCAGCAAGGTGGGGC 3000  
3232 AAGCAGATGCTCCAGGATTTTACACCTGAGCCCTGCCACCCCTGCTGAAAACAT 3291  
3001 AAGCAGATGCTCCAGGATTTTACACCTGAGCCCTGCCACCCCTGCTGAAAACAT 3060  
3292 CCGCAGCTGAAGAGACAGAGGAGGTGGCAGAGTTTACCTGGGGAACACAAACAGGGA 3351  
3061 CCGCAGCTGAAGAGACAGAGGAGGTGGCAGAGTTTACCTGGGGAACACAAACAGGGA 3120  
3352 TCTTTTCTGCCCTGCTCCAGTCCAGTGGCCTGA 3387  
3121 TCTTTTCTGCCCTGCTCCAGTGGCCTGA 3156

RESULT 8  
AAV71603  
ID AAV71603 standard; cDNA; 3156 BP.  
XX  
AC AAV71603;  
XX  
DT 04-FEB-1999 (first entry)  
XX  
DE Human NF-KB-inducing kinase (NIK) polypeptide encoding cDNA.  
XX  
KW NIK; IKB; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;  
tumour necrosis factor; binding; genetic hybridisation; screening;  
signal transduction; biopharmaceutical; human; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2844  
FT /\*tag= a  
FT /product= "NIK polypeptide"  
XX  
XX US5843721-A.  
XX  
XX 01-DEC-1998.  
XX  
XX 03-JUL-1997; 97US-00887518.  
XX  
XX 03-JUL-1997; 97US-00887518.  
XX  
XX (TULA-) TULARIK INC.  
XX  
XX Wu L, Rothe M;  
XX  
XX WPI: 1999-044580/04.  
XX  
XX P-PSDB; AAW81564.  
XX  
XX Probe, vector or recombinant nucleic acid encoding a polypeptide, useful  
especially human nuclear factor kappa-B-inducing kinase protein - useful  
for producing recombinant protein.  
XX  
XX Claim 1; Col 11-14; 15pp; English.  
XX  
XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)  
polypeptide. The NIK polypeptide or its fragment has one or more  
activities selected from kinase activity and inhibitory activity; IKB  
kinase-alpha and beta binding activity and binding inhibitory activity;  
tumour necrosis factor (TNF) receptor-associated factor 2 binding  
activity and binding inhibitory activity; IKB binding activity and  
binding inhibitory activity. NF-kB activating and inhibitory activity. A  
vector containing the NIK nucleic acid can be used to transform host

CC cells for the recombinant production of the protein. The NIK nucleic acid  
CC and the polypeptide may be used in diagnosis (e.g. genetic hybridisation  
CC screen for NIK transcripts), therapy (e.g. NIK kinase inhibitors to  
CC inhibit TNF signal transduction), and in the biopharmaceutical industry  
XX  
SQ Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;

Query Match 68.5%; Score 3149.6; DB 2; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 232 ATGGCAGTGATGGAATGGCTGCCAGTGGCCCTGGCTCAGCAGTGGGCGCAGCAGAG 291  
DB 1 ATGGCAGTGATGGAATGGCTGCCAGTGGCCCTGGCTCAGCAGTGGGCGCAGCAGAG 60  
QY 292 GAACTCCCAAGCAAGAGAGAGAGCGCGCCACCTGGGGAGAAACACAGAGCTCCGCTTAC 351  
DB 61 GAACTCCCAAGCAAGAGAGAGAGCGCGCCACCTGGGGAGAAACACAGAGCTCCGCTTAC 120  
QY 352 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 411  
DB 121 AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 180  
QY 412 GTGATTACCAAGGGCAGACCAAGGAAGGCTCCGAGGAGGCGCAGCTGCCATCTCTATC 471  
DB 181 GTGATTACCAAGGGCAGACCAAGGAAGGCTCCGAGGAGGCGCAGCTGCCATCTCTATC 240  
QY 472 ATCGCCAGGCTGAGTGTGAGAAATAGCAAGATTCAGGCCACCTTTTCAGAACGCATT 531  
DB 241 ATCGCCAGGCTGAGTGTGAGAAATAGCAAGATTCAGGCCACCTTTTCAGAACGCATT 300  
QY 532 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 591  
DB 301 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 360  
QY 592 GTGGCCCATGTACAGAGGGCAAAATGGCCCGTGTGTTGGAGAGGAAAGGTCGCGAGC 651  
DB 361 GTGGCCCATGTACAGAGGGCAAAATGGCCCGTGTGTTGGAGAGGAAAGGTCGCGAGC 420  
QY 652 AAAGCCGGAGAAACGGAAGAGAGTCAAAATGCCCTGGCTCATGCAAGGAGTGGCC 711  
DB 421 AAAGCCGGAGAAACGGAAGAGAGTCAAAATGCCCTGGCTCATGCAAGGAGTGGCC 480  
QY 712 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGAGAGCTGCACCATCCCACTGAGGAG 771  
DB 481 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGAGAGCTGCACCATCCCACTGAGGAG 540  
QY 772 GATGAGTCTCCACTCGGCGCCCATATGTTTAAAGAACACCCCGCAGTTTCAACAGCCTCTG 831  
DB 541 GATGAGTCTCCACTCGGCGCCCATATGTTTAAAGAACACCCCGCAGTTTCAACAGCCTCTG 600  
QY 832 AAGGAACAGGCTTGGGCACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGCT 891  
DB 601 AAGGAACAGGCTTGGGCACTCTGTTTAAAGCAGCTTGGGAGGCGCTACGCGCGCT 660  
QY 892 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGGCAATGTCTGAACCAAGTGGG 951  
DB 661 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTGGCAATGTCTGAACCAAGTGGG 720  
QY 952 AAATGCAACCAACCCAGGAGCGGCGCCCTGTCCTGCGCCACGACCCCTTCCCTTAT 1011  
DB 721 AAATGCAACCAACCCAGGAGCGGCGCCCTGTCCTGCGCCACGACCCCTTCCCTTAT 780  
QY 1012 AGCAGATGCTCATCCCTTCCCATTCACCCCTCTCAGGCGCTGGAACCTCAGCCTCTG 1071  
DB 781 AGCAGATGCTCATCCCTTCCCATTCACCCCTCTCAGGCGCTGGAACCTCAGCCTCTG 840  
QY 1072 GAGTCTTCTTGGGCAAACTGGCTGTGTAGACAGCAGAAACCTTGCCTGAGCCACAC 1131  
DB 841 GAGTCTTCTTGGGCAAACTGGCTGTGTAGACAGCAGAAACCTTGCCTGAGCCACAC 900  
QY 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGGCC 1191  
DB 1191 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGGCC 2040

DB 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGGCC 960  
QY 1192 AGCTGCTGTCTCGTGTGGTCCCATCAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1251  
DB 961 AGCTGCTGTCTCGTGTGGTCCCATCAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1020  
QY 1252 CTGCAAGGCGAGCTGAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGGCCAAAGCCTGG 1311  
DB 1021 CTGCAAGGCGAGCTGAGCTCAAGCCAGGCCCAAGCCTGACAGCCTGGCCAAAGCCTGG 1080  
QY 1312 GCAGCAGGGGCTCCAGATCCCGGAGGCCAGCCCAAACTGAGGACAAAGAGGGTGTCT 1371  
DB 1081 GCAGCAGGGGCTCCAGATCCCGGAGGCCAGCCCAAACTGAGGACAAAGAGGGTGTCT 1140  
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCC 1431  
DB 1141 CTGCTCACTGAGAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCC 1200  
QY 1432 ACACAAGCTCCGCTGGGCGAGAGGCTCTTTCGGAGAGGTGCACAGATGGAGCAAG 1491  
DB 1201 ACACAAGCTCCGCTGGGCGAGAGGCTCTTTCGGAGAGGTGCACAGATGGAGCAAG 1260  
QY 1492 CAGACTGGCTCCAGTGGGCTGTCAAAAGGTGGCCCTGGAAGTATTTTCGGGCGAGAGG 1551  
DB 1261 CAGACTGGCTCCAGTGGGCTGTCAAAAGGTGGCCCTGGAAGTATTTTCGGGCGAGAGG 1320  
QY 1552 CTGATGGCATGTGAGGATTCAGCTCACCAGAAATTTGTCCTTTGTATGGAGCTGTGAGA 1611  
DB 1321 CTGATGGCATGTGAGGATTCAGCTCACCAGAAATTTGTCCTTTGTATGGAGCTGTGAGA 1380  
QY 1612 GAAGGGCTTTGGGTCAAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671  
DB 1381 GAAGGGCTTTGGGTCAAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440  
QY 1672 GTCAAGGAGCAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTG 1731  
DB 1441 GTCAAGGAGCAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTG 1500  
QY 1732 GAGGGTCTGGAATACCTCCACTCAGAAAGGATTCGATGGGGAGCTCAAAAGCTGACAAAC 1791  
DB 1501 GAGGGTCTGGAATACCTCCACTCAGAAAGGATTCGATGGGGAGCTCAAAAGCTGACAAAC 1560  
QY 1792 GTGCTCTGTCTCAGCGATGGGAGCCACGACCCCTCTGTGTGATTTTGGCCATGTGTGTGT 1851  
DB 1561 GTGCTCTGTCTCAGCGATGGGAGCCACGACCCCTCTGTGTGATTTTGGCCATGTGTGTGT 1620  
QY 1852 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGGAGGACTACATCCCTGGGCGACAGAG 1911  
DB 1621 CTTCAACCTGATGGCTGGGAAAGTCTTGTCTCAGGAGGACTACATCCCTGGGCGACAGAG 1680  
QY 1912 ACCACATGGCTCCGAGAGTGGTGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGG 1971  
DB 1681 ACCACATGGCTCCGAGAGTGGTGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGG 1740  
QY 1972 AGCAGCTGTGTATGATGTGCAATGCTCAACGGCTGGCCACCCCTGGACTCAGTTCCTTC 2031  
DB 1741 AGCAGCTGTGTATGATGTGCAATGCTCAACGGCTGGCCACCCCTGGACTCAGTTCCTTC 1800  
QY 2032 CGAGGGCGCTCTGCTCCTCAAGATTTGCGAGGAGCTCCGCTGTGAGGAGATCCACCC 2091  
DB 1801 CGAGGGCGCTCTGCTCCTCAAGATTTGCGAGGAGCTCCGCTGTGAGGAGATCCACCC 1860  
QY 2092 TCCTGCGCCCTCTCAGAGCCAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCAC 2151  
DB 1861 TCCTGCGCCCTCTCAGAGCCAGGCCATCAAGAGGGGCTGAGGAAAGGCCCATCCAC 1920  
QY 2152 CGCCTGTCTGCGAGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGT 2211  
DB 1921 CGCCTGTCTGCGAGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGT 1980  
QY 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2271  
DB 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2040

QY 2272 AATTACACAGACCTCCATCCCGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2331  
DB 2041 AATTACACAGACCTCCATCCCGCCGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2100  
QY 2332 CGGCCAGCTGAGGAGCAACAGCGCAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAG 2391  
DB 2101 CGGCCAGCTGAGGAGCAACAGCGCAGAGCCCTTAAGCTCCAGCTCTCTCCACACAGAG 2160  
QY 2392 CCCCAGAGCCAAAAGAGTCTCTCCCTTGACTTTGAGCAAGAGAGTCTGGGATGTGG 2451  
DB 2161 CCCCAGAGCCAAAAGAGTCTCTCCCTTGACTTTGAGCAAGAGAGTCTGGGATGTGG 2220  
QY 2452 GAACCTTTACCTCTGCTCTCCCTGAGCCAGCCCTTGCAGAAACCCAGAGTCCACAGAG 2511  
DB 2221 GAACCTTTACCTCTGCTCTCCCTGAGCCAGCCCTTGCAGAAACCCAGAGTCCACAGAG 2280  
QY 2512 CGGAAAGCAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTTCTCTCAAC 2571  
DB 2281 CGGAAAGCAACCGTCCCGAGCAGGAACCTGCAGCAGCTGGAATAGAAATTTCTCTCAAC 2340  
QY 2572 AGCTGTCCAGCCATTTCTCTGAGAGAGAGGAGCAAAATTTCTCTGTGCTCTCAGATC 2631  
DB 2341 AGCTGTCCAGCCATTTCTCTGAGAGAGAGGAGCAAAATTTCTCTGTGCTCTCAGATC 2400  
QY 2632 GACAGCTCTCCCTGTCGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCG 2691  
DB 2401 GACAGCTCTCCCTGTCGATGACAGTGAGAGAACCCATCAAGGCCCTCTCAAGCTCG 2460  
QY 2692 CGGGACACCTCAGCTCAGGCGTACACTCTCGAGCAGCCAGCGAGGCTCGAAGCTCC 2751  
DB 2461 CGGGACACCTCAGCTCAGGCGTACACTCTCGAGCAGCCAGCGAGGCTCGAAGCTCC 2520  
QY 2752 AGCTGGAACATGCTGTGCTCCCGGGCGGGCCGACCGACACCCCAAGCTATTTCAATGTT 2811  
DB 2521 AGCTGGAACATGCTGTGCTCCCGGGCGGGCCGACCGACACCCCAAGCTATTTCAATGTT 2580  
QY 2812 GTGAAGTCAAAATACAGTCTTAAATGTTGACACCTGCACATCGGGAGTTCCACCGG 2871  
DB 2581 GTGAAGTCAAAATACAGTCTTAAATGTTGACACCTGCACATCGGGAGTTCCACCGG 2640  
QY 2872 GTCAAAAGTGGAGACATCGCCATCGGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2931  
DB 2641 GTCAAAAGTGGAGACATCGCCATCGGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGC 2700  
QY 2932 TTGGTCAACAAAGACGGGAGCCTGTTGCTGATGAGAGTGCAGACTCGGGGCATC 2991  
DB 2701 TTGGTCAACAAAGACGGGAGCCTGTTGCTGATGAGAGTGCAGACTCGGGGCATC 2760  
QY 2992 GACCTGAGTGACACTGGCCCTGATGCGAGCTTCGCTGGAGCTGGAGGTTCAAGCAT 3051  
DB 2761 GACCTGAGTGACACTGGCCCTGATGCGAGCTTCGCTGGAGCTGGAGGTTCAAGCAT 2820  
QY 3052 GGCAGCTGGAGAAAGGGCCCTAACCTGCTCCCTCCACCCGGCTCCACACTGCCGAAA 3111  
DB 2821 GGCAGCTGGAGAAAGGGCCCTAACCTGCTCCCTCCACCCGGCTCCACACTGCCGAAA 2880  
QY 3112 GCAGCCTTCTGCTCGGTGCAAGATGCTGCCTTGAAAAACACAGGCTCAGCGTTCCAGG 3171  
DB 2881 GCAGCCTTCTGCTCGGTGCAAGATGCTGCCTTGAAAAACACAGGCTCAGCGTTCCAGG 2940  
QY 3172 GGAATGGCAGCCCGGCTCAGAGTGGNAACAGGGCTTCGAGCAGCAAGTGGGGC 3231  
DB 2941 GGAATGGCAGCCCGGCTCAGAGTGGNAACAGGGCTTCGAGCAGCAAGTGGGGC 3000  
QY 3232 AAGCAGATGCTCCAGGATTTTCACTTGAGCCCTGCCCACTGCTGTAAGAAACAT 3291  
DB 3001 AAGCAGATGCTCCAGGATTTTCACTTGAGCCCTGCCCACTGCTGTAAGAAACAT 3060  
QY 3292 CCGCCACGTGAAGAGACAGAGGAGATGCGAGATTTACCTGGGGAACAAAGACAGGA 3351  
DB 3061 CCGCCACGTGAAGAGACAGAGGAGATGCGAGATTTACCTGGGGAACAAAGACAGGA 3120

QY 3352 TCTTTTCTGCCCCCTGCTCCAGTCCAGTTCAGTTCGCTTGA 3387  
DB 3121 TCTTTTCTGCCCCCTGCTCCAGTCCAGTTCAGTTCGCTTGA 3156  
RESULT 9  
AAV69285  
ID AAV69285 standard; cDNA; 3156 BP.  
XX AAV69285;  
AC AAV69285;  
XX 04-FEB-1999 (first entry)  
XX Human NF-kB-inducing kinase (NIK) polypeptide encoding cDNA.  
DE NIK; Ikb; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;  
KW tumour necrosis factor; binding; genetic hybridisation; screening;  
KW signal transduction; biopharmaceutical; immunogen; pharmacological;  
KW transcription regulator; human; ds.  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT CDS 1..2844  
FT /tag= a  
FT /product= "NIK polypeptide"  
FT  
XX US844073-A.  
XX 01-DEC-1998.  
XX 13-FEB-1998; 98US-00023321.  
XX 03-JUL-1997; 97US-00887518.  
XX (TULA-) TULARIK INC.  
XX Wu L, Rothe M;  
XX WPI; 1999-044664/04.  
XX P-PSDB; AAW81561.  
XX New isolated peptide comprising a specified 947 amino acid sequence - has  
PT e.g. kinase activity, kinase inhibitory activity, Ikb kinase-alpha  
PT binding activity, and Ikb kinase-alpha binding inhibitory activity.  
XX Disclosure; Col 11-14; 15pp; English.  
XX This cDNA encodes a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)  
CC polypeptide. The NIK polypeptide or its fragment has one or more  
CC activities selected from kinase activity and inhibitory activity; Ikb  
CC kinase-alpha and beta binding activity and binding inhibitory activity;  
CC tumour necrosis factor (TNF) receptor-associated factor 2 binding  
CC activity and binding inhibitory activity; Ikb binding activity and  
CC binding inhibitory activity, NF-kB activating and inhibitory activity.  
CC The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g.  
CC genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK  
CC kinase inhibitors to inhibit TNF signal transduction), and in the  
CC biopharmaceutical industry (e.g. as immunogens, reagents for isolating  
CC other transcription regulators, and reagents for screening chemical  
CC libraries for pharmacological agents)  
XX  
SQ Sequence 3156 BP; 736 A; 991 C; 895 G; 534 T; 0 U; 0 Other;  
Query Match 68.5%; Score 3149.6; DB 2; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 232 ATGGCAGTGCATGAATGGCTGCGCCAGTCCCTGGCTCAGAGTGGGGCAGAGAAG 291  
DB 1 ATGGCAGTGCATGAATGGCTGCGCCAGTCCCTGGCTCAGAGTGGGGCAGAGAAG 60  
QY 292 GAACTCCCCCAAGCCAAAGAGGAGAACGCGCCACTGGGGAGAGAAACAGAGCTCCGCTTAC 351









Db 361 GTGGCCCATGCTACAGAGGGGAAAATGGCCGTGTGTGTTGGAAAGGAAAGCGTGCAGC 420  
Qy 652 AAAGCCCGAAGAAACGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCAAGGAGTGGCC 711  
Db 421 AAAGCCCGAAGAAACGGAAGAAAGAGCTCAAAAGTCCCTGGCTCATGCAAGGAGTGGCC 480  
Qy 712 TTGGCCAAACCCCTCCCGAGACCCCTGAGCAGAGAGAGTGCACATCCCAAGTGCAGGAG 771  
Db 481 TTGGCCAAACCCCTCCCGAGACCCCTGAGCAGAGAGAGTGCACATCCCAAGTGCAGGAG 540  
Qy 772 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 831  
Db 541 GATGAGTCTCCACTCGGCGCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 600  
Qy 832 AAGGAACACAGGCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGGCTACGGCGGGCT 891  
Db 601 AAGGAACACAGGCTTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGGCTACGGCGGGCT 660  
Qy 892 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGTCGAAACCAAGTGTGG 951  
Db 661 CTGCTCGATCAGAACTCCAAACTGATCAGCCCTTTGCAATGTCGAAACCAAGTGTGG 720  
Qy 952 AAACTGCACACCCCGAGAGCGAGGCCCCCTGCCCCCTGCCACGCAACCCCTTCCCTAT 1011  
Db 721 AAACTGCACACCCCGAGAGCGAGGCCCCCTGCCCCCTGCCACGCAACCCCTTCCCTAT 780  
Qy 1012 AGCAGACTGCTCATCCCTTCCATTCACCCCTCTCCAGCCCTGGAACCTCACCCTCTG 1071  
Db 781 AGCAGACTGCTCATCCCTTCCATTCACCCCTCTCCAGCCCTGGAACCTCACCCTCTG 840  
Qy 1072 GAGTCTTCTCGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGAGCCACAC 1131  
Db 841 GAGTCTTCTCGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGAGCCACAC 900  
Qy 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCGCTGGCCCAACCTGAGGCC 1191  
Db 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCGCTGGCCCAACCTGAGGCC 960  
Qy 1192 AGCTGCTGTCTCGTGTGGCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1251  
Db 961 AGCTGCTGTCTCGTGTGGCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCT 1020  
Qy 1252 CTGCAAGGCAGCTGAGCTCAAGCCAGGCCACACAGCTGACAGCTTGGCCAAAGACCTGG 1311  
Db 1021 CTGCAAGGCAGCTGAGCTCAAGCCAGGCCACACAGCTGACAGCTTGGCCAAAGACCTGG 1080  
Qy 1312 GCAGCAGGGGCTCCAGATCCCGGAGCCCGAGCCCAAACTGAGGACAAAGAGGTGTC 1371  
Db 1081 GCAGCAGGGGCTCCAGATCCCGGAGCCCGAGCCCAAACTGAGGACAAAGAGGTGTC 1140  
Qy 1372 CTGCTCACTGAGAACTCAAGCCAGTGAATATGAGTACCGAGAGAGTCCACTGGGCC 1431  
Db 1141 CTGCTCACTGAGAACTCAAGCCAGTGAATATGAGTACCGAGAGAGTCCACTGGGCC 1200  
Qy 1432 ACGCACCAGCTCCGCTCGGCGAGAGGCTCTTTCGAGAGGTGCACAGGATGGAGACAAG 1491  
Db 1201 ACGCACCAGCTCCGCTCGGCGAGAGGCTCTTTCGAGAGGTGCACAGGATGGAGACAAG 1260  
Qy 1492 CAGACTGGCTTCCAGTGGCTGTCAAAAGGTGGCCCTGGAAATTTTCGGGCGAGAGGAG 1551  
Db 1261 CAGACTGGCTTCCAGTGGCTGTCAAAAGGTGGCCCTGGAAATTTTCGGGCGAGAGGAG 1320  
Qy 1552 CTGATGGCATGTGAGGATGACCTCACCAGAAATGTTCCTTTGTATGAGAGCTGTGAGA 1611  
Db 1321 CTGATGGCATGTGAGGATGACCTCACCAGAAATGTTCCTTTGTATGAGAGCTGTGAGA 1380  
Qy 1612 GAAGGGCTTGGTCAACATCTTCAAGAGTGTGGAAGTGGCTCCCTGGGCCAGCTG 1671  
Db 1381 GAAGGGCTTGGTCAACATCTTCAAGAGTGTGGAAGTGGCTCCCTGGGCCAGCTG 1440  
Qy 1672 GTCAAGGAGCAGGGTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGGCCCTG 1731

Db 1441 GTCAGGAGCAGGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCGAGGCCCTG 1500  
Qy 1732 GAGGCTCTGGAATACCTCACTCAAGAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAAC 1791  
Db 1501 GAGGCTCTGGAATACCTCACTCAAGAGGATTTCTGCATGGGGAGCTCAAAAGCTGACAAC 1560  
Qy 1792 GTGCTCTGTCCAGGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851  
Db 1561 GTGCTCTGTCCAGGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620  
Qy 1852 CTTCAACTGATGCTGGGAAAGTCTTGTCTCAAGGGGACTCATCCCTGGGACACAG 1911  
Db 1621 CTTCAACTGATGCTGGGAAAGTCTTGTCTCAAGGGGACTCATCCCTGGGACACAG 1680  
Qy 1912 ACCACATGGCTCCGAGAGTGGTGTCTGGGACAGGAGCTGCGACGCAAGGTGATGCTGG 1971  
Db 1681 ACCACATGGCTCCGAGGTTGTCTGGGACAGGAGTGCAGCGCAAGGTGATGCTGG 1740  
Qy 1972 AGCAGCTGTGTATGATGTGCAATGCTCAAGGGTGCACCCCTGGACTCAGTTCTTC 2031  
Db 1741 AGCAGCTGTGTATGATGTGCAATGCTCAAGGGTGCACCCCTGGACTCAGTTCTTC 1800  
Qy 2032 CGAGGGCGCTCTGCCTCAAGATTTGCCAGGAGCTCCGCTGTGAGGAGATCCACCC 2091  
Db 1801 CGAGGGCGCTCTGCCTCAAGATTTGCCAGGAGCTCCGCTGTGAGGAGATCCACCC 1860  
Qy 2092 TCCTCGCGCCCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCATCCAC 2151  
Db 1861 TCCTCGCGCCCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGCCATCCAC 1920  
Qy 2152 CGCTGTCTGAGCGGAGTGGGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGT 2211  
Db 1921 CGCTGTCTGAGCGGAGTGGGAGGAAGTGAACCGGGCACTACAGCAAGTGGAGGT 1980  
Qy 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2271  
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2040  
Qy 2272 AATTACCAACAGACCCCTCCATCCAGCCAGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2331  
Db 2041 AATTACCAACAGACCCCTCCATCCAGCCAGAGAGAGCTTTCCGCAAGGGCCCCAGGGCCC 2100  
Qy 2332 CGGCCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTTCCAGCTCTCTCTCCACACAG 2391  
Db 2101 CGGCCAGCTGAGGAGACAACAGGAGAGCCCTTAAGCTTCCAGCTCTCTCTCCACACAG 2160  
Qy 2392 CCCCAGAGCCAAACAAGTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGG 2451  
Db 2161 CCCCAGAGCCAAACAAGTCTCTCTCTGACTTTGAGCAAGGAGAGTCTGGGATGTGG 2220  
Qy 2452 GAACCCCTTACTCTGTCTCTCCCTGGAGCCAGCCCTTCGAGAAACCCAGCTCACCAGAG 2511  
Db 2221 GAACCCCTTACTCTGTCTCTCTGGAGCCAGCCCTTCGAGAAACCCAGCTCACCAGAG 2280  
Qy 2512 CGGAAAGCAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAAATATGAAATATTCTCAAC 2571  
Db 2281 CGGAAAGCAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAAATATGAAATATTCTCAAC 2340  
Qy 2572 AGCTGTCCAGCCATTTTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATC 2631  
Db 2341 AGCTGTCCAGCCATTTTCTGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATC 2400  
Qy 2632 GACAGCTCTCTCTGTGAGTGAAGAGAACCCATCAAGGCTCTCAAGCTCG 2691  
Db 2401 GACAGCTCTCTCTGTGAGTGAAGAGAACCCATCAAGGCTCTCAAGCTCG 2460  
Qy 2692 CGGACACCTCTGAGCTCAGGGGTACATCTCTGAGCAGCCAGGGCGAGGCTCGAAGCTCC 2751  
Db 2461 CGGACACCTCTGAGCTCAGGGGTACATCTCTGAGCAGCCAGGGCGAGGCTCGAAGCTCC 2520  
Qy 2752 AGCTGGAACATGTGTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2811  
Db 2521 AGCTGGAACATGTGTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580

QY 2812 GTGAAAGTCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGAGTTCCACCGG 2871  
DB |||||  
QY 2581 GTGAAAGTCAAATACAGTCTCTTAATGTTGAACACCTGCACATCCGGAGTTCCACCGG 2640  
DB |||||  
QY 2872 GTCAAAGTGGGAGACATGCCACTGGCATCAGACGCCAGATCCCAGCTGCAGCCTTCAGC 2931  
DB |||||  
QY 2641 GTCAAAGTGGGAGACATGCCACTGGCATCAGACGCCAGATCCCAGCTGCAGCCTTCAGC 2700  
QY 2932 TTGGTCAACCAAGACGGGAGCCTGTTGCTACGACATGGAGTGCAGACTCGGGCATC 2991  
DB |||||  
QY 2701 TTGGTCAACCAAGACGGGAGCCTGTTGCTACGACATGGAGTGCAGACTCGGGCATC 2760  
QY 2992 GACCTGCAGTGACACTGGCCCTGATGGCAGCTTCGCCTGAGCTGGAGGGTCAAGCAT 3051  
DB |||||  
QY 2761 GACCTGCAGTGACACTGGCCCTGATGGCAGCTTCGCCTGAGCTGGAGGGTCAAGCAT 2820  
QY 3052 GCCCAGCTGGAGAACAGGCCCTAA 3075  
DB |||||  
QY 2821 GCCCAGCTGGAGAACAGGCCCTAA 2844

## RESULT 11

ADQ67410  
ID ADQ67410 standard; cdNA; 3152 BP.

AC ADQ67410;

XX DT 07-OCT-2004 (first entry)

DE XX Novel human cdNA sequence #2383.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;  
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;  
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
KW cancer.

XX OS Homo sapiens.

XX PN EP1440981-A2.

XX PD 28-JUL-2004.

XX PF 21-JAN-2004; 2004EP-00001196.

XX PR 21-JAN-2003; 2003JP-00102206.

XX PR 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

DR P-PSDB; ADQ67717.

XX Novel 2495 cdNA, useful for treating osteoporosis, neurological diseases,

PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

PS Claim 1; SEQ ID NO 4571; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded  
CC polypeptides, sequences hybridizing to these nucleotides, sequences  
CC encoding partial polypeptides and sequences having 70% or 90% identity to  
CC the nucleotide and protein sequences. The nucleotides and polypeptides  
CC are useful as diagnostic markers or therapeutic target for the diseases  
CC or morbid states. They are also useful for treating osteoporosis,  
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,  
CC dementia and various cancers. This sequence corresponds to a nucleotide  
CC sequence of the invention.

XX Sequence 3152 BP; 704 A; 1000 C; 874 G; 574 T; 0 U; 0 Other;

Query Match 58.2%; Score 2673.4; DB 12; Length 3152;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 36; Indels 10; Gaps 7;

QY 1764 TCTGATGGGACGTCAAAGCTGACACGTTGCTTCAACCTGATGCCCTGGGAAAGTCTCTTGT 1823  
DB |||||  
QY 1824 CCTCTGTCACTTTTGGCCATGCTGTGTCTTCAACCTGATGCCCTGGGAAAGTCTCTTGT 1883  
DB |||||  
QY 400 CCTCTGTCACTTTTGGCCATGCTGTGTCTTCAACCTGATGCCCTGGGAAAGTCTCTTGT 459  
QY 1884 CACAGGGGACTACATCTCCCTGGCACAGACCCACATGGCTCCGGAGTGTCTGGGCGAG 1943  
DB |||||  
QY 460 CACAGGGGACTACATCTCCCTGGCACAGACCCACATGGCTCCGGAGTGTCTGGGCGAG 519  
QY 1944 GAGCTGCCAGCGCAAGGTGGATGTCTGGAGCAGCTGTGTATGATGTGTCACATGCTCAA 2003  
DB |||||  
QY 520 GAGCTGCCAGCGCAAGGTGGATGTCTGGAGCAGCTGTGTATGATGTGTCACATGCTCAA 579  
QY 2004 CGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCTGCCCTCAAGATTGCCAGCGA 2063  
DB |||||  
QY 580 CGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCTGCCCTCAAGATTGCCAGCGA 639  
QY 2064 GCCTCCGCTGTGAGGGAGATCCCAACCTCTCTCGGCCCTCTCTCACAGCCCGAGCCATCCA 2123  
DB |||||  
QY 640 GCCTCCGCTGTGAGGGAGATCCCAACCTCTCTCGGCCCTCTCTCACAGCCCGAGCCATCCA 699  
QY 2124 AGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGGAGGGAGCTGGGAGGGAAGGT 2183  
DB |||||  
QY 700 AGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGGAGGGAGCTGGGAGGGAAGGT 759  
QY 2184 GAACCGGGCACTACAGCAAGTGGGAGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGA 2243  
DB |||||  
QY 760 GAACCGGGCACTACAGCAAGTGGGAGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGA 819  
QY 2244 ACCAAGACATCCACCGCAAAATCAAGCCAAATTAACACAGAGCCCTCCATGCCAGCCGAG 2303  
DB |||||  
QY 820 ACCAAGACATCCACCGCAAAATCAAGCCAAATTAACACAGAGCCCTCCATGCCAGCCGAG 879  
QY 2304 AGAGCTTTGCGCAAGGGCCCAAGGGCCCGGCGCAGCTGAGGAGACAACAGGGAGAGCCCC 2363  
DB |||||  
QY 880 AGAGCTTTGCGCAAGGGCCCGGCGCAGCTGAGGAGACAACAGGGAGAGCCCC 939  
QY 2364 TAAGCTCCAGCTCTCTCCACAGAGCCCCCAGAGCAACAAAGTCTCTCTCCCTTGAC 2423  
DB |||||  
QY 940 TAAGCTCCAGCTCTCTCTCCACAGAGCCCCCAGAGCAACAAAGTCTCTCTCCCTTGAC 999  
QY 2424 TTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCCCTTACCTCTGTCTCTCCCTGGAGCCAGC 2483  
DB |||||  
QY 1000 TTTGAGCAAGGAGGAGTCTGGGATGTGGAAACCCCTTACCTCTGTCTCTCCCTGGAGCCAGC 1059  
QY 2484 CCTTCCAGAAACCCAGCTACAGAGGGGAAAGCAACCGTCCCGGAGCAGAACTGCA 2543  
DB |||||  
QY 1060 CCTTCCAGAAACCCAGCTACAGAGGGGAAAGCAACCGTCCCGGAGCAGAACTGCA 1119  
QY 2544 GCAGCTGGAAATAGAAATTTCTCAACAGCCCTGTCCAGCCATTTTCTCTGAGAGACA 2603  
DB |||||  
QY 1120 GCAGCTGGAAATAGAAATTTCTCAACAGCCCTGTCCAGCCATTTTCTCTGAGAGACA 1179  
QY 2604 GGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCCCTCTCCCTGTGCGATGACAGTGAAGA 2663  
DB |||||  
QY 1180 GGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCCCTCTCCCTGTGCGATGACAGTGAAGA 1239  
QY 2664 GAACCCATCAAGGGCTCTCAAGCTCGGGAGACACCTGAGCTCAGGGGTACACTCTTG 2723  
DB |||||  
QY 1240 GAACCCATCAAGGGCTCTCAAGCTCGGGAGACACCTGAGCTCAGGGGTACACTCTTG 1299  
QY 2724 GAGCAGCCAGGCCGAGGCTCGAAAGCTCCAGCTGGAAACATGGTGTGGCCCCGGGGCGGC 2783  
DB |||||  
QY 1300 GAGCAGCCAGGCCGAGGCTCGAAAGCTCCAGCTGGAAACATGGTGTGGCCCCGGGGCGGC 1359  
QY 2784 CACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAAATACAGTCTCTTTAATGGTGA 2843

```
Db 1360 CACCGACACCCCAAGCTATTTCAATGGTGTGAAGTCCAAATACAGTCTCTTAATGGTGA 1419
Qy 2844 ACACCTGCACATCCCGGAGTCCACCGGGTCAAAAGTGGGAGACATCGGCCACTGGGATCAG 2903
Db 1420 ACACCTGCACATCCCGGAGTCCACCGGGTCAAAAGTGGGAGACATCGGCCACTGGGATCAG 1479
Qy 2904 CAGCCAGATCCAGCTGAGCTTCCAGCTTGGTTCACAAAGACGGGAGCCTGTTCGCTA 2963
Db 1480 CAGCCAGATCCAGCTGAGCTTCCAGCTTGGTTCACAAAGACGGGAGCCTGTTCGCTA 1539
Qy 2964 CGACATGAGGTGCGAGCTCGGATCGAGCTGAGTGCACATCGGCCCTGATGGCAG 3023
Db 1540 CGACATGAGGTGCGAGCTCGGATCGAGCTGAGTGCACATCGGCCCTGATGGCAG 1599
Qy 3024 CTTGGCTGAGCTGAGGAGTCAAGCATGGGAGCTGGAGAACAGGCCCTTAACCTTGCCC 3083
Db 1600 CTTGGCTGAGCTGAGGAGTCAAGCATGGGAGCTGGAGAACAGGCCCTTAACCTTGCCC 1659
Qy 3084 TCCACCGCGGCTCCACACTGCGGAAAGAGCCTTCTGCTGCGGTGCACATGCTGCC 3143
Db 1660 TCCACCGCGGCTCCACACTGCGG--AAGCAGCCTTCTGCTGCGGTGCACATGCTGCC 1718
Qy 3144 TGAACAACAGGCTCAGCGCTTCCAGGGATTCGACGCCCGCGGCTCAGTGGGAAC 3203
Db 1719 TG-AAACAACAGGCTCAGCGCTTCCAGGGATTCGAC--CCCCGGGCTCAGAGTGGGAC 1776
Qy 3204 CAGGCGCTCGAGCAGCAGAGTGGGGCAAGCAGAAATGCTCCAGGATTTTCAACCTGA 3263
Db 1777 CAGGCGCTCGAGCAGCAGAGTGGGGCAAGCAGAAATGCTCCAGGATTTTCAACCTGA 1836
Qy 3264 GCCCTGCCCCCCTGCTGAAACAA-TCGCGCACGTGAAGACAGAGGAGATGGC 3322
Db 1837 GCCCTGCCCCCCTGCTGAGAAACACTCGCCACAGTGAAGAGACAG-AGGAGGATGGC 1895
Qy 3323 AGGAGTTACCTGGGGAAACAAACAGGGATCTTTTCTGCCCCCTGCTCCAGTCCAGTTGG 3382
Db 1896 AGGAGTTACCTGGGGAAACAAACAGG---ATCTTCTGCCCCCTGCTCCAGTCCAGTTGG 1951
Qy 3383 CTTGACCCGCTTGGATCAGTCAACATTTGTTGGCAGACAGGGAGAGCAGCTTCCAGCCT 3442
Db 1952 CTTGACCCGCTTGGATCAGTCAACATTTGTTGGCAGACAGGGAGAGCAGCTTCCAGCCT 2011
Qy 3443 GGGTCAGAGGGGTGGGAGCCCTTCGCGCCCTCACCCTCCAGGCTGCTGTGAGAGTGT 3502
Db 2012 GGGTCAGAGGGGTGGGAGCCCTTCGCGCCCTCACCCTCCAGGCTGCTGTGAGAGTGT 2071
Qy 3503 CAAGTGTGAAGGGCCCAAACTCAGGTTCAAGTGCAGAACAGGTCAGAGGATGCCCGC 3562
Db 2072 CAAGTGTGAAGGGCCCAAACTCAGGTTCAAGTGCAGAACAGGTCAGAGGATGCCCGC 2131
Qy 3563 CCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTCACCTGGCCAGCTCACCCCT 3622
Db 2132 CCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTCACCTGGCCAGCTCACCCCT 2191
Qy 3623 TTTGGGTGTAGGGGAAAAAGATGCTGACCTCGGGAAGGCTCCCTGGTAGAATACACCAC 3682
Db 2192 TTTGGGTGTAGGGGAAAAAGATGCTGACCTGGGAAGGCTCCCTGGTAGAATACACCAC 2251
Qy 3683 ACTTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGGTTTCAGCCCAAGGACCAAA 3742
Db 2252 ACTTTTTCAGGTTGTGCAACACAGGTCCTGAGTTGACCTCTGGTTTCAGCCCAAGGACCAAA 2311
Qy 3743 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCGCAGACATGTCCTTTGCTGGTGA 3802
Db 2312 GAAGGTGTGAAGTGAAGTGTCTCAGTCCCGCAGACATGTCCTTTGCTGGTGA 2371
Qy 3803 CCACCTCTTCCAGAGCAGAGGCCCCGAGCCCTTTCAGGGCCAGCAGCTGCCCCAGACTC 3862
Db 2372 CCACCTCTTCCAGAGCAGAGGCCCCGAGCCCTTTCAGGGCCAGCAGCTGCCCCAGACTC 2431
Qy 3863 GCTGCGACTCAGTTCCCTCATCTGTAAAGGTGAAGGGGTGATCGCAGGATATCCCTGACAGG 3922
```

```
Db 2432 GCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTATGACAGGATATGCTGACAGG 2491
Qy 3923 AACAGTCTGTGGATGGACATGATCAGTGTCTAAGAAAGACAGCAGAGAGAGACG-TCGGGC 3981
Db 2492 AACAGTCTGTGGATGGACATGATCAGTGTCTAAGAAAGACAGCAGAGAGAGACGTCGGC 2551
Qy 3982 GCCCCAGCCCCACTATCAGTGTCCAGGTGCTGTTCCCGACAGACAGCTCAGCATCAC 4041
Db 2552 GCCCCAGCCCCACTATCAGTGTCCAGGTGCTGTTCCCGACAGACAGCTCAGCATCAC 2611
Qy 4042 ACTCACACTACCTGCGCTGCCCTGGCCACAGAGGGTACTGCCACAGGCACTTTGCACTC 4101
Db 2612 ACTCACACTACCTGCGCTGCCCTGGCCACAGAGGGTACTGCCACAGGCACTTTGCACTC 2671
Qy 4102 TGATGACCTCAAGACACTTTTCATGGCTGCCCTTGGCAGGGCAGGGCAGGCACTGACAC 4161
Db 2672 TGATGACCTCAAGACACTTTTCATGGCTGCCCTTGGCAGGGCAGGGCAGGCACTGACAC 2731
Qy 4162 TGTAGGACATAGCAAGCCAGGAGATGGGTGAAGGGACACAGTCTTTGAGCTGTCCACAT 4221
Db 2732 TGTAGGACATAGCAAGCCAGGAGATGGGTGAAGGGACACAGTCTTTGAGCTGTCCACAT 2791
Qy 4222 GCATGTCACTCTCAAACTCTTCCAGATTTCTTAAGAAATAGCACCCCTTCCCAATG 4281
Db 2792 GCATGTCACTCTCAAACTCTTCCAGATTTCTTAAGAAATAGCACCCCTTCCCAATG 2851
Qy 4282 CCCAGCTTACGCTTCTTCCAGGGAGCTTCTCAGGAGCTCAGCTAGCATTAATCAGC 4341
Db 2852 CCCAGCTTACGCTTCTTCCAGGGAGCTTCTCAGGAGCTCAGCTAGCATTAATCAGC 2911
Qy 4342 TGTGAATCGTCAGGGGTGCTGTAGCTCAACCTCCTGGGGCAGGGGAGCCGAGACT 4401
Db 2912 TGTGAATCGTCAGGGGTGCTGTAGCTCAACCTCCTGGGGCAGGGGAGCCGAGACT 2971
Qy 4402 CCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGCCTTTGTCAGCTGTCCACAT 4461
Db 2972 CCGTGGGAGAGCTCATTTCCACATCTTGCCAGACAGCCTTTGTCAGCTGTCCACAT 3031
Qy 4462 GAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCCGACGACATAAAGAACTGAGCCCT 4521
Db 3032 GAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCCGACGACATAAAGAACTGAGCCCT 3091
Qy 4522 GGTACTCGAGAGTCTGGTTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGA 4581
Db 3092 GGTACTCGAGAGTCTGGTTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGGTGATGA 3151
Qy 4582 C 4582
Db 3152 C 3152

RESULT 12
ACN44437
ID ACN44437 standard; cDNA; 2828 BP.
XX
XX ACN44437;
XX
XX 18-NOV-2004 (first entry)
XX
XX Mouse mRNA sequence mC117748.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.
XX
XX Mus musculus.
XX
XX WO2003073826-A2.
XX
XX 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX
XX
```

(SAGR-) SAGRES DISCOVERY.

Morris DW:

WPI; 2003-328604/31.

Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

comprises a nucleotide sequence.

Claim 1; SEQ ID NO 884; Opp; English.

The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA gene carcinoma. Note: This patent is an equivalent to basic patent US20020182586A1, for which no sequence data was published.

The present invention relates to novel DNA and protein sequences which

are associated with carcinomas. The sequences are useful for: (i) for

screening drug candidates; (ii) for screening of bioactive agent capable

of binding to Carcinoma Associated Protein (CAP); (iii) for screening of

a bioactive agent capable of modulating the activity of CAP; (iv) for

evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating

carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

(x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining carcinoma associated (CA) gene count number. In addition, the

determining Carcinoma Associated (CA) gene copy number. In addition, the

CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA gene.

carcinoma including lymphoma. The present sequence is one such CA gene sequence. Note: This patent is an equivalent to basic patent.

sequence. NOTE: this patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

032002182388AT, FOR WHICH NO SEQUENCE DATA WAS PUBLISHED

Sequence 2828 BP; 688 A; 880 C; 790 G; 470 T; 0 U; 0 Other;

Query Match 44.4%; Score 2042.6; DB 11; Length 2828;

Best Local Similarity 83.4%; Pred. No. 0;

Matches 2377; Conservative 0; Mismatches 444; Indels 28; Gaps 4;

232 ATGGCAGTGATGGAAATGGCCCTGCCACAGTGCCCTGGCTCAGCAGTGGGCAGCAGAAG 291

\_\_\_\_\_

292 GAACTCCCCAAGCCAAGGAGAGACGCCGCGCCACTGGGGAAGAAACAGAGCTCCGTCTAC 351

1	_____
2	_____
3	_____
4	_____
5	_____
6	_____
7	_____
8	_____
9	_____
10	_____
11	_____
12	_____
13	_____
14	_____
15	_____
16	_____
17	_____
18	_____
19	_____
20	_____
21	_____
22	_____
23	_____
24	_____
25	_____
26	_____
27	_____
28	_____
29	_____
30	_____
31	_____
32	_____
33	_____
34	_____
35	_____
36	_____
37	_____
38	_____
39	_____
40	_____
41	_____
42	_____
43	_____
44	_____
45	_____
46	_____
47	_____
48	_____
49	_____
50	_____
51	_____
52	_____
53	_____
54	_____
55	_____
56	_____
57	_____
58	_____
59	_____
60	_____
61	_____
62	_____
63	_____
64	_____
65	_____
66	_____
67	_____
68	_____
69	_____
70	_____
71	_____
72	_____
73	_____
74	_____
75	_____
76	_____
77	_____
78	_____
79	_____
80	_____
81	_____
82	_____
83	_____
84	_____
85	_____
86	_____
87	_____
88	_____
89	_____
90	_____
91	_____
92	_____
93	_____
94	_____
95	_____
96	_____
97	_____
98	_____
99	_____
100	_____

352 AAGCTTGAGGCCCGTGGAGAAGAGGCCCTGTGTCTGCGGAAAGTGGGAGATCCTGAATGAC 411

5' ATGCTTGGAGCCCGCCTGTGCTGCGAATCCGATGAC 411

412 GTGATTA CCA GGGCA CAGCC A CCGA AGGCT CCGC CAGCC CAGCT CCGT CCGT C

412 GIGATACCAAGGGCACHAGCCAAAGGAAGGCATCCGAGGCAGGGCAGCTGGCATCTCTATC 471

4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 10

472 ATCGCCAGGCTGAGTGTGAGATAGCCAAAGAGTTCAGCCCCACCTTTTCAGAACGCATT 531

532 TTCAATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT 591

.....

592 GTGGCCCATGCTACAGAGGGCAAAATGCCCGTGTGTGTGGAAGGAAAGCGTCGCAGC 651

420

652 AAAGCCCGAAGAAACGGAAGAAGAGAGCTCAAAGTCCCTGGCTCATGCAGGAGTGGCC 711

421 AATAAGCCCGAATTAACCTAAGGAAGAAAGAGTCTCAATAGTCACTGGCCCAAGGCAGGAGTGGCC 480

712 TTGGCCAAACCCCTCCCCAGGACCCCTGAGCAGGAGAGCTGCACCATCCCAGTGCAGGAG 771

481 TTAGCCAAGCCCCCTGGCCAGAAACCCCTGAGCAAGAGAGAGCTGTACCATCCAGTACAGGAA 540

772 GATGAGTCTCCACTCGGGGGCCCCATATGTTAGAAACACCCCGCAGTTCA CCAAGCCTCTTG 831

541 GATGAGTCTCCACTAGGCAACCTCTATGCCAGAAATGTCTCCCCAGTTCACCAAGCCTCTG 600

Db 1580 ACGAGACCCACATGGCCACCAAGAGTGGTGTATGGAAAGCCCTCGATGCCAAGTGGAC 1739  
Qy 1566 GTCTGGAGACAGCTCTGTATGATGCTGCACATGCTCAACGGCTCCACCCCTGGACTCAG 2025  
Db 1740 ATCTGGAGCAGCTGTGATGATGCTCCACATGCTCAAGGGCTGCCACCCCTGGACTCAG 1799  
Qy 2026 TTCTTCGGAGGGCCGCTCTGCTCAAGATTGCCAGCGAGCTCGCCCTGTGAGGGAGATC 2085  
Db 1800 TACTTCGGAGGGCCGCTTTGTCTCAAGATTGCCAGCGAGCTCCACCGATCAGGGAGATC 1859  
Qy 2086 CCACCTCTCTCGGCCCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGGCC 2145  
Db 1860 CCACCTCTCTCGGCCCTCTCACAGCCAGGCCATCAAGAGGGGCTGAGGAAAGAGGCC 1919  
Qy 2146 ATCCACCGCTGTCTGACGCGAGCTGGAGGGAGGTGAACCGGGCACTACAGCAAGTG 2205  
Db 1920 GTCCACCGAGCATCTGCCATGGAGCTTCGGAGGAAAGTGGCAAGGCACACAGCAAGTG 1979  
Qy 2206 GGAGGTCTGAAGAGCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACGCCCAAT 2265  
Db 1980 GGAGGTCTGAAGAGCCCTTGGAGGAGGAATATAAAGAACCAAGACCTCCACGCCCAAG 2039  
Qy 2266 CAAGCCCAATTACCAACAGACCTCTCATGCCAGCCGAGAGAGCTTCGCCAAGGGCCCCA 2325  
Db 2040 CAAGCCCACTGCCACAGACCTTACTCTCCGCGAGAGA----- 2080  
Qy 2326 GGGCCCGGCGAGCTGAGGAGACCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCCA 2385  
Db 2081 --GAACCCACAGCCAGGCCCAACAGAGCGGGCTCTGAGCCTCAGCCTCTCTACCG 2138  
Qy 2386 CCAGAGCCCCAGAGCCCAAGATCTCTCTCCCTTGACTTTGAGCAAGAGGAGTCTGGG 2445  
Db 2139 CCAGAACCCAGAACCCGAGCAAGCGCCAGCCCTGAACCTGAGCAAGGAGGAGTCTGGC 2198  
Qy 2446 ATGTGGGAACCTTACTCTCTCTCTCTGAGCGAGCCCTCCAGAAACCCAGCTCA 2505  
Db 2199 ACATGGGAACCCCTGCTCTCTCTCTGAGCCAGCCACTGCCAAGGCCCGAGCTC 2258  
Qy 2506 CCAGAGGGGAAGCAACCGTCCCGGAGCAGGAATGCGAGAGCTGGAAATAGATTATTC 2565  
Db 2259 CCAGACCGGAGGCAACCTTGCAGAGCTGGAGCTACAGCAACTGGAGATAGACTGTT 2318  
Qy 2566 CTCAACAGCTGTCCAGCCATTTCTCTGAGGAGCAGGAGCAAAATCTCTGTGCTC 2625  
Db 2319 CTCAACAGCTGTCCAGCCGTTCTCTGTGAGGAACAGGAACAAATCTCTCTGCTC 2378  
Qy 2626 AGCATGCAGACCTCTCCCTGTGCGATCAGAGTGAAGACCCATCAAGGCCCTCTCAA 2685  
Db 2379 AGCATGCAGACCTCTCCCTGTGCGATCAGAGTGAAGAAATCATCGAAGGCCCTCTCAG 2438  
Qy 2686 AGCTCGGGGACACCTTGAGCTCAGGGGTACACTCTCTGGAGCAGCGCGAGGCTCGA 2745  
Db 2439 AGCTCAGGGACACCTTGAGTTCTGGGTGACCTTTGGAAACAGCCAGCTGAGCAAGA 2498  
Qy 2746 AGCTCAGCTGGAATAGTGTCTGGCCGGGGGGCCCAACCGACACCCCAAGCTATTTC 2805  
Db 2499 ACCTGAGCTGACAGCGGCTGGCCGGGGGGCGCTACTGACATCCCGAGCTACTTC 2558  
Qy 2806 AATGGTGTGAAGTCCAAATACAGTCTTAAATGGTGAACACCTGCACATCCGGAGTTC 2865  
Db 2559 AACGGGTCAAGTCCAGATCCAGTCTCTCAATGGCGAAACACCTGCAATATCCGGGAATTC 2618  
Qy 2866 CACCGGTCAAGTGGGAGACATCGCACTGGCATCAGCAGCCAGATCCAGCTGCAGCC 2925  
Db 2619 CACCGGTCAAGTGGGAGACATTCACCGGCAATCAGCAGCCAGATCCAGCCACAGCT 2678  
Qy 2926 TTCAAGTGTGTACCAAGAGCGGAGCCTGTTCGCTACGACATGAGAGGTGCCAGACTCG 2985  
Db 2679 TTCAAGCTGTGTACCAAGAGTGGACAGCCTGTTTGTATGATGGAGGTGCCAGACTCG 2738  
Qy 2986 GGCATCGACCTGAGTGCACACTGGCCCTCTGATGGCAGCTTCGGCTGAGAGTGGAGGTC 3045  
Db 2739 GGCATCGACCTGAGTGCACCTGGCCCTCTGATGGCAGCTTTGTGTCGACCTGGAGGTC 2798

Qy 3046 AAGCATGGCCAGCTGGAGAAACAGGCCCTTA 3074  
Db 2799 AAGCATGGTCAGCTGGAGAACCGACCTTA 2827

## RESULT 13

AAV03324  
ID AAV03324 standard; cDNA; 2631 BP.

XX AAV03324;

XX 15-APR-1998 (first entry)

XX Clone 10 cDNA encoding NMPL1, a TRAF2 binding protein.

XX Human tumour necrosis factor receptor-associated factor 2; TRAF2;  
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;  
KW intracellular signalling activity; acute hepatitis;  
KW autoimmune-induced cell death; ss.

XX Homo sapiens.

XX W09737016-A1.

XX 09-OCT-1997.

XX 01-APR-1997; 97WO-IL000117.

XX 02-APR-1996; 96IL-00117800.

XX 26-AUG-1996; 96IL-00119133.

XX (VEDA ) YEDA RES & DEV CO LTD.

XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

XX NPI; 1997-503101/46.

XX DNA encoding tumour necrosis factor receptor-associated factor binding  
PT molecule - used for modulation or mediation in cells of the activity of  
PT NF-kB.

XX Claim 4; Fig 4; 127pp; English.

XX The present sequence encodes a protein designated NMPL1, a TRAF2 binding  
CC protein. This protein contains Ser/Thr protein kinase motifs. A full  
CC length cDNA clone encoding NF-kappaB (NIK) was obtained by PCR using the  
CC present clone, clone 10. The clone 10 protein is capable of binding to at  
CC least amino acids 222-501 of TRAF2. The TRAF-2 binding proteins can be  
CC used for modulation or mediation in cells of the activity of NF-kappaB or  
CC any other intracellular signalling activity modulated or mediated by  
CC TRAF2. TRAF-binding proteins are especially used for prevention or  
CC treatment of pathological conditions associated with NF-kB induction,  
CC e.g. acute hepatitis, autoimmune-induced cell death, e.g. death of the  
CC beta langerhans cells or the pancreas that results in diabetes, the death  
CC of cells in graft rejection, the death of oligodendrocytes in the brain  
CC in multiple sclerosis, and AIDS-inhibited T cell suicide which causes  
CC proliferation of the AIDS virus and hence the AIDS disease. The proteins  
CC are also useful for screening of ligands capable of binding to a protein,  
CC which are useful for modulating cellular activity modulated/mediated by  
CC TRAF2

SQ Sequence 2631 BP; 596 A; 791 C; 704 G; 459 T; 0 U; 81 Other;

Query Match 40.3%; Score 1851.2; DB 2; Length 2631;  
Best Local Similarity 90.5%; Pred. No. 0;  
Matches 2335; Conservative 4; Mismatches 159; Indels 81; Gaps 34;

Qy 2099 CCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 2158

Db 1 CCCCTCTCACAGCCAGGCCATCCAGAGGGGCTGAGAAAGAGCCCATCCACCGCGTGT 60

Qy 2159 CTGACGCGAGCTGGGAGGGAAGTGAACCCGGGCATCTACAGCAAGTGGAGGTCTGAAGA 2218



Qy	4301	CCGAGGGAGCTA-CTCAGGACCTCACGTAGCATTTAAATCAGCTGTG-AAATCGTCAGGGG	4358
Db	2281	CCGAGGGAGCTA-CTCAGGACCTCACGTAGCATTTAAATCAGCTGTGNAATCGTCAGGGG	2340
Qy	4359	TGTCTGTAGCCTCAACTCTCTGGGCGAGGGACGCCGAGACTCCGTGGGAGAGCTCAT	4418
Db	2341	TGTCTGTAGCCTCAACTCTCTGGGCGAGGGACGCCGAGACTCCGTGGGAGAGCTCAT	2400
Qy	4419	TCCACACATCTTGCCAAAGACAGCCCTTT--GTCCAGCTGTGCCATTTGAGTCAGAGCTGCTCCC	4477
Db	2401	TCCACACATCTTGCCAAAGACAGCCCTTTGTCAGCTGTCCACATTTGAGTCAGAGCTGCTCCC	2460
Qy	4478	GGGAGAGAGCCGGCCCGCCAGACATAAAGACTCGACGCTTGGTACTCGACAGCTCG	4537
Db	2461	GGGAGAGAGCCGGCCCGCCAGACATAAAGAACTCGACGCTTGGTACTCGACAGCTCG	2520
Qy	4538	GGTTGTAGAGAACTCTTTGTGAAGCAATATAAGTTTGGGGTGATGACAAATCTTAAAAAA	4596
Db	2521	GGTTGTAGAGAACTCTTTGTGAAGCAATATAAGTTTGGGGTGATGACAAATCTTAAAAAA	2579

## RESULT 14

ACN44438  
ID ACN44438 standard; DNA: 73967 BP.

18-NOV-2004 (first entry)

XX  
DE Human genomic sequence hCG27607.

XX  
KW  
Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX  
OS Homo sapiens.

XX PN WO2003073826-A2.

XX PD 12-SEP-2003.

XX  
PF 28-FEB-2003: 2003WO-US006235.XX  
PR 01-MAR-2002: 2002US-00087192.XX  
PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW:

XX  
DR WPI: 2003-328604/31.

xx PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.

XX  
PS  
Claim 1: SEQ ID NO 886: Opp: English.

The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biotrap; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Sequence 73967 BP; 18655 A; 18476 C; 19039 G; 17797 T; 0 U; 0 Other;

Query Match 33.9%; Score 1556.2; DB 11; Length 73967;

Best Local Similarity 98.0%; Pred. No. 0; Matches 1650; Conservative 0; Mismatches 23; Indels 10; Gaps 7;										
QY	2909	AGATCCAGCTGTCAGCCTTTCAGCTTGGTTCACCAAGAGCGGCGAGCCTGTTGGCTACGACA	2968							
DB	62243	AGATCCAGCTGTCAGCCTTTCAGCTTGGTTCACCAAGAGCGGCGAGCCTGTTGGCTACGACA	62302							
QY	2969	TGAGGTGTCAGACTCGGGCATCGACTCGAGTGCACACTGGCCCTGANTGGCAGCTTCG	3028							
DB	62303	TGAGGTGTCAGACTCGGGCATCGACTCGAGTGCACACTGGCCCTGANTGGCAGCTTCG	62362							
QY	3029	CCTGGAGCTGGAGGTCAGAGCATGCCAGTGGAGAACAGCGCCCTAACCCCTGCCCTCCAC	3088							
DB	62363	CCTGGAGCTGGAGGTCAGAGCATGCCAGTGGAGAACAGCGCCCTAACCCCTGCCCTCCAC	62422							
QY	3089	CGCGGCTCCACACTGCCGGAAGCAGCCTTCTGCTCGGTGCACGATGCTGCCCTGAAA	3148							
DB	62423	CGCGGCTCCACACTGCCGGAAGCAGCCTTCTGCTCGGTGCACGATGCTGCCCTGAAA	62480							
QY	3149	ACAGGCTCAGCGGTTCCAGGGGATTGCCAGCCCCCGGCTCACAGTGGGAACAGGG	3208							
DB	62481	ACAGGCTCAGCGGTTCCAGGGGATTGCCAGCCCCCGGCTCACAGTGGGAACAGGG	62539							
QY	3209	CCTCGCAGCAGCAAGGTGGGGCAGCAGATGCTCTCCAGGATTTACACTGAGCCCT	3268							
DB	62540	CCTCGCAGCAGCAAGGTGGGGCAGCAGATGCTCTCCAGGATTTACACTGAGCCCT	62599							
QY	3269	GCCCACTCTGCTGAAAAACA--TCGGCCAGCTGAAGAGAGACAGAAAGGAGTGGCAGGAG	3327							
DB	62600	GCCCACTCTGCTGAAAAACA--TCGGCCAGCTGAAGAGAGAGTGGCAGGAG	62658							
QY	3328	TTACTCTGGGAAACAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTGCAGTGGCCCTGA	3387							
DB	62659	TTACTCTGGGAAACAAAACAGGATCTTTTCTGCCCCCTGCTCCAGTGCAGTGGCCCTGA	62714							
QY	3388	CCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGGGAGAGCAGCTTCCAGCCTGGGTC	3447							
DB	62715	CCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGGGAGAGCAGCTTCCAGCCTGGGTC	62774							
QY	3448	AGAAGGGTGGCGAGCCCTTCGGCCCCCTCACCTTCAGGCTGCTGTCAGAGTGTCAAAGT	3507							
DB	62775	AGAAGGGTGGCGAGCCCTTCGGCCCCCTCACCTTCAGGCTGCTGTCAGAGTGTCAAAGT	62834							
QY	3508	GTGTAAAGGCCCAAACTCAGGTTTCAGTGCAGAACAGGTCAGCAGGATGCCCCGCCGTA	3567							
DB	62835	GTGTAAAGGCCCAAACTCAGGTTTCAGTGCAGAACAGGTCAGCAGGATGCCCCGCCGTA	62894							
QY	3568	GGTTAAGGGGGCCCTCTAAACCCCTTGGCTGGCCCTCACCTGGCCAGCTCACCCCTTTGG	3627							
DB	62895	GGTTAAGGGGGCCCTCTAAACCCCTTGGCTGGCCCTCACCTGGCCAGCTCACCCCTTTGG	62954							
QY	3628	GTGTAGGGGAAAAGAAATGCTCACCTTCGGGAAGGCTCCTGCTAGAATACACCACTTT	3687							
DB	62955	GTGTAGGGGAAAAGAAATGCTCACCTTCGGGAAGGCTCCTGCTAGAATACACCACTTT	63014							
QY	3688	TCAGGTTGTTGCAACACAGGTCCTCAGTTGACCTCTGCTTCAGCCAAAGGACCAAGAGG	3747							
DB	63015	TCAGGTTGTTGCAACACAGGTCCTCAGTTGACCTCTGCTTCAGCCAAAGGACCAAGAGG	63074							
QY	3748	TGTGTAAGTGAAAGTGTCTCAGTCCCAGACATGTGCCCTTTGCTGCTGGCTACCACT	3807							
DB	63075	TGTGTAAGTGAAAGTGTCTCAGTCCCAGACATGTGCCCTTTGCTGCTGGCTACCACT	63134							
QY	3808	CTTCCCAGCAGCAGCGCCCGAGCCCTTCAGGCCCAGCAGCTGCCCCAGACTCGCTGG	3867							
DB	63135	CTTCCCAGCAGCAGCGCCCGAGCCCTTCAGGCCCAGCAGCTGCCCCAGACTCGCTGG	63194							
QY	3868	CACCTAGTTTCCCTCATCTGTAAAGTGGAAGGTGATGCAGGATATGCTGCAGAGAAACAG	3927							
DB	63195	CACCTAGTTTCCCTCATCTGTAAAGTGGAAGGTGATGCAGGATATGCTGCAGAGAAACAG	63254							
QY	3928	TCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAGCAGAGAGACG--TCCGGCCGCC	3986							



Db 63255 TCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAGCAGAGAGACGCTCGCGGCGCCC 63314  
QY 3987 AGCCCCACTATCAGTGTCCAGCGTGTGTTCCCCAGAGACAGAGCTCAGCATCACACTGA 4046  
Db 63315 AGCCCCACTATCAGTGTCCAGCGTGTGTTCCCCAGAGACAGAGCTCAGCATCACACTGA 63374  
QY 4047 CACTCACCTGCCCCTGCCCTGGCCAGAGGGTACTGCCAGCGCACTTTTGCACTCTGTATG 4106  
Db 63375 CACTCACCTGCCCCTGCCCTGGCCAGAGGGTACTGCCAGCGCACTTTTGCACTCTGTATG 63434  
QY 4107 ACTCTAAACACTTTTCATGGCTGCCCTCTGGCAGGCGAGGCGAGGCACTGTATG 4166  
Db 63435 ACTCTAAACACTTTTCATGGCTGCCCTCTGGCAGGCGAGGCGAGGCACTGTATG 63494  
QY 4167 GAGCATAGCAAGCAGGAGATGGGTGAAGGAGACAGTCTTGAGCTGTCCACATGCAATG 4226  
Db 63495 GAGCATAGCAAGCAGGAGATGGGTGAAGGAGACAGTCTTGAGCTGTCCACATGCAATG 63554  
QY 4227 TGACTCTCTCAAACTCTTCCAGATTCTCTAAGAATAGCACCCCTTCCCAATGGCCCA 4286  
Db 63555 TGACTCTCTCAAACTCTTCCAGATTCTCTAAGAATAGCACCCCTTCCCAATGGCCCA 63614  
QY 4287 GCTTAGCTCTTCTCCAGGAGACTACTCAGGATCAGGATCAATTAATCAGCTGTGA 4346  
Db 63615 GCTTAGCTCTTCTCCAGGAGACTACTCAGGATCAGGATCAATTAATCAGCTGTGA 63674  
QY 4347 ATCGTCAGGGGTGTCTGTAGCTCAACCTCTCGGGCAGGCGAGCTCCGTCG 4406  
Db 63675 ATCGTCAGGGGTGTCTGTAGCTCAACCTCTCGGGCAGGCGAGCTCCGTCG 63734  
QY 4407 GGAGAGCTCATTTCCACATCTTGCCAGACAGACCTTTGTCCAGCTGTCCACATTTAGTTC 4466  
Db 63735 GGAGAGCTCATTTCCACATCTTGCCAGACAGACCTTTGTCCAGCTGTCCACATTTAGTTC 63794  
QY 4467 AGACTGCTCCCGGGAGAGAGCCCGCGCCCGCCAGCACATAAAGAACTGCGAGCTTGGTATC 4526  
Db 63795 AGACTGCTCCCGGGAGAGAGCCCGCGCCCGCCAGCACATAAAGAACTGCGAGCTTGGTATC 63854  
QY 4527 TCGAGAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGTATGACAAAT 4586  
Db 63855 TCGAGAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAATAAAGTTTGGGTGTATGACAAAT 63914  
QY 4587 GTT 4589  
Db 63915 GTT 63917

RESULT 15  
AAS79491/c  
ID AAS79491 standard; cDNA; 2760 BP.  
XX  
AC AAS79491;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #15295.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX

PI Drmanac RT, Liu C, Tang YT;  
XX WPI: 2001-639362/73.  
DR P-PSDB; ABG15304.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 15295; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2760 BP; 486 A; 840 C; 810 G; 624 T; 0 U; 0 Other;  
Query Match 32.1%; Score 1474.4; DB 5; Length 2760;  
Best Local Similarity 73.4%; Pred. No. 0;  
Matches 2370; Conservative 0; Mismatches 6; Indels 852; Gaps 4;

QY 232 ATGGCAGTGTGAAATGGCTTCCAGGTGCCCTGAGTGGCTGAGTGGGCGAGCAGAG 291  
Db 2760 ATGGCAGTGTGAAATGGCTTCCAGGTGCCCTGAGTGGCTGAGTGGGCGAGCAGAG 2701  
QY 292 GAACTCCCAAGCCAAAGGAGAGAGCGCCACTTGGGGAAGAAACAGAGCTCCGTCTTAC 351  
Db 2700 GAACTCCCAAGCCAAAGGAGAGAGCGCCACTTGGGGAAGAAACAGAGCTCCGTCTTAC 2641  
QY 352 AAGCTTTAGGCGCGTGGAGAGAGCCCTGTGTCTTCCGAAAAGTGGGAGATCCTGAATGAC 411  
Db 2640 AAGCTTTAGGCGCGTGGAGAGAGCCCTGTGTCTTCCGAAAAGTGGGAGATCCTGAATGAC 2581  
QY 412 GTGATTACCAAGGCGCAGCCAGAGAGGCTCCGAGGCGAGGCGCGAGTCCCATCTCTATC 471  
Db 2580 GTGATTACCAAGGCGCAGCCAGAGAGGCTCCGAGGCGAGGCGCGAGTCCCATCTCTATC 2521  
QY 472 ATCGCCCAAGGCTGAGT----- 487  
Db 2520 ATCGCCCAAGGCTGAGTGTAGGGGCTTAGGGTGTGTTGGGCACGGGGCGAGGCGCTTACC 2461  
QY 488 ----- 487  
Db 2460 ATCTTCCACTAGAGTGGGCCCAACCCCAAGGAATAGCACCGTGTCTTTGAGCGCTGCTG 2401  
QY 488 -----GTGAAATAGCAAGAGTTTTCAGCCCCCACC 516  
Db 2400 GGCCAGTCTTGGGCGGCGAGGCTGTGCAGGTTGAGAAATAGCAAGAGTTTTCAGCCCCCACC 2341  
QY 517 TTTTCAGAAACGATTTTTCATCGCTGGGTCCAAAACAGATACAGCCAGTCCGAGAGTCTTGAT 576  
Db 2340 TTTTCAGAAACGATTTTTCATCGCTGGGTCCAAAACAGATACAGCCAGTCCGAGAGTCTTGAT 2281  
QY 577 CAGATCCCAACAATGTGGCCCATGCTTACAGAGGGCAAAATGGGCCCGGTGTGTGTTGGAAG 636



Db 2280 CAGATCCCAACATGTGGCCATCTACAGAGGCAAAATGGCCCGTGTGTGGGAAG 2221  
Qy 637 GGAAGCGTCGAGCAAAAGCCGGAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCT 696  
Db 2220 GGAAGCGTCGAGCAAAAGCCGGAGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCT 2161  
Qy 697 CATGAGAGTGGCTTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAGCTGCACC 756  
Db 2160 CATGAGAGTGGCTTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAGCTGCACC 2101  
Qy 757 ATCCAGTGCAGGAGGATGAGTCTCCACTCGCGCGCCCATATGTTAGAAAACACCCCGCAG 816  
Db 2100 ATCCAGTG----- 2092  
Qy 817 TTCACCAAGCCCTGAAGGAACACGAGCCTTGGGCAACTCTGTTTAAAGCAGCTGGCGAG 876  
Db 2091 ----- 2092  
Qy 877 GGCCTACGGCCGGCTCTGCCCTGATCAGAACTCCACAAACTGATCAGCCCCCTTGCATGT 936  
Db 2091 ----- 2092  
Qy 937 CTGAACACGTTGGAACATGCACACCCCGAGGAGGAGGCCCTGCCCTGCCACG 996  
Db 2091 ----- 2092  
Qy 997 CACCCCTTCCCTATAGCAGACTGCCCTCATCCCTTCCCATTTCCACCCCTCTCCAGCCCTGG 1056  
Db 2091 -----CAGCCCTGG 2083  
Qy 1057 AAACCTACCTCTGAGTCTCTTCTGGGCAAACTGGCCTGTGTAGACAGCAAGAAACC 1116  
Db 2082 AAACCTACCTCTGAGTCTCTTCTGGGCAAACTGGCCTGTGTAGACAGCAAGAAACC 2023  
Qy 1117 TTGCTGACCCACACCTGAGCAAACTGCCCTGTGTAGACAGTCCAAAGCCCTGCCTGGC 1176  
Db 2022 TTGCTGACCCACACCTGAGCAAACTGCCCTGTGTAGACAGTCCAAAGCCCTGCCTGGC 1963  
Qy 1177 CCACACCTGGAGCCAGCTGCTGTCTGTGGTCCCATGAGAAAGTTTCTGTGGAGAA 1236  
Db 1962 CCACACCTGGAGCCAGCTGCTGTCTGTGGTCCCATGAGAAAGTTTCTGTGGAGAA 1903  
Qy 1237 TACCTAGTGCATGCTCTCAAGGACAGCTGAGCTCAAGCCAGGCCACAGCCTGACCAGC 1296  
Db 1902 TACCTAGTGCATGCTCTCAAGGACAGCTGAGCTCAGGCCAGGCCACAGCCTGACCAGC 1843  
Qy 1297 CTGGCCAGACTGGGACAGCGGGCTCCAGATCCCGGGAGCCCGAGCCCAAAACTGAG 1356  
Db 1842 CTGGCCAAAGACTGGGACAGCGGGCTCCAGATCCCGGGAGCCCGAGCCCAAAACTGAG 1783  
Qy 1357 GACAAACGAGGGTCTCTCTCACTGAGAAACTCAAGCCAGTGGATTATGATCCCGAGAA 1416  
Db 1782 GACAAACGAGGGTCTCTCTCACTGAGAAACTCAAGCCAGTGGATTATGATCCCGAGAA 1723  
Qy 1417 GAAGTCACTGGGCCACGACACAGCTCCGCTGGGACAGAGCTCTTCTGGAGAGGTGCAC 1476  
Db 1722 GAAGTCACTGGGCCACGACACAGCTCCGCTGGGACAGAGCTCTTCTGGAGAGGTGCAC 1663  
Qy 1477 AGGATGGAGGACAGCAGCTGGCTTCCAGTGGCTGTCAAAA----- 1519  
Db 1662 AGGATGGAGGACAGCAGCTGGCTTCCAGTGGCTGTCAAAAAGGAGCAGTCGCGAGGGTG 1603  
Qy 1520 ----- 1519  
Db 1602 GCCGTGGGTCTGACTTGTGTGAACAGAACTTGGCTATCACTAGCCACTGAGGTCCGTGGA 1543  
Qy 1520 ----- 1519  
Db 1542 ACTGMAAGGAGCAGCCTGTCACTGTCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGTA 1483  
Qy 1520 ----- 1519

Db 1482 TGTGAGGCGAGTGTGAGAGGACACACGTGGCCTCGGTGTCTCAATACAGCCAGGGCCTG 1423  
Qy 1520 -----AGGTGGC 1527  
Db 1422 AGGGCAAGTCCCAACGTGGAAGATGTTGGAGCGTCAATGAAATTCACACTGAGGTGCG 1363  
Qy 1528 CTGGAAGTATTTCCGGCAGAGGAGCTGATGGCATGTGCAGATTTGACCTCACCCAGAAAT 1587  
Db 1362 CTGGAAGTATTTCCGGCAGAGGAGCTGATGGCATGTGCAGATTTGACCTCACCCAGAAAT 1303  
Qy 1588 GTCCCTTTGTATGAGCTGTGAGAGAAAGGGCTTTGGGTCAACATCTTCATCGAGCTGCTG 1647  
Db 1302 GTCCCTTTGTATGAGCTGTGAGAGAAAGGGCTTTGGGTCAACATCTTCATCGAGCTGCTG 1243  
Qy 1648 GAAGGTGGCTCCCTGGGCCAGCTGTCAAGGAGCAGGGCTGTCTCCAGAGAGACCGGGCC 1707  
Db 1242 GAAGGTGGCTCCCTGGGCCAGCTGTCAAGGAGCAGGGCTGTCTCCAGAGAGACCGGGCC 1183  
Qy 1708 CTGTACTACTGGGCCAGGCCCTGGAGGCTGTGAANTACTTCCAATCAGGAAGATTTCTG 1767  
Db 1182 CTGTACTACTGGGCCAGGCCCTGGAGGCTGTGAANTACTTCCAATCAGGAAGATTTCTG 1123  
Qy 1768 CATGGGACGTCGAAGCTGCAACGTGCTCTGTCCAGCGATGGAGCCACGCGAGCCCTC 1827  
Db 1122 CATGGGACGTCGAAGCTGCAACGTGCTCTGTCCAGCGATGGAGCCACGCGAGCCCTC 1063  
Qy 1828 TGTGACTTTGGCCATGTGTGTCTTTCAACCTGATGGCCTGGGAAAGTCTTGTCTACA 1887  
Db 1062 TGTGACTTTGGCCATGTGTGTCTTTCAACCTGATGGCCTGGGAAAGTCTTGTCTACA 1003  
Qy 1888 GGGACTATATCTCTGSCACAGAGACCCACATGGCTCCGGAGTGTGTCTGGGACAGGAGC 1947  
Db 1002 GGGACTATATCTCTGSCACAGAGACCCACATGGCTCCGGAGTGTGTCTGGGACAGGAGC 943  
Qy 1948 TGGACCCCAAGTGGATGTCTGGAGCAGTGTGTATGATGCTGCATGCTCAACGGC 2007  
Db 942 TGGACCCCAAGTGGATGTCTGGAGCAGTGTGTATGATGCTGCATGCTCAACGGC 883  
Qy 2008 TGCACACCTCGGACTCAGTTCTTCCGAGGGCCGCTCTGCCCTCAAGATTGGCCAGGAGCCT 2067  
Db 882 TGCACACCTCGGACTCAGTTCTTCCGAGGGCCGCTCTGCCCTCAAGATTGGCCAGGAGCCT 823  
Qy 2068 CCGCTGTGAGGAGATCCCACTCTGTGGCCCTCTCAAGCCCGAGGCTCAAGAG 2127  
Db 822 CCGCTGTGAGGAGATCCCACTCTGTGGCCCTCTCAAGCCCGAGGCTCAAGAG 763  
Qy 2128 GGGCTGAGGAAGAGCCATCCACCGCTGTCTGAGGGAGCTGGGAGGAGGTGAC 2187  
Db 762 GGGCTGAGGAAGAGCCATCCACCGCTGTCTGAGCGGAGCTGGGAGGAGGTGAC 703  
Qy 2188 CGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCA 2247  
Db 702 CGGCACTACAGCA----- 689  
Qy 2248 AGACATCCAACCGCAATCAAGCAATTAACAAGAGCCCTTCAATGCCCCAGAGAG 2307  
Db 688 ----- 689  
Qy 2308 CTTTGGCAAGGGCCCCCAGGGCCCCCGGCGAGCTGAGGAGACACAGGAGAGCCCTAAG 2367  
Db 688 ----- 689  
Qy 2368 CTCAGGCTCTCTCTCCACAGAGCCCGGAGCAAAAGTCTCTCTCTTGCATTTG 2427  
Db 688 -----AGAGCCCCCAGGCCAAACAAAGTCTCTCTCTTGCATTTG 649  
Qy 2428 AGCAAGGAGGAGTGTGGATGTGGAAACCTTACCTGTCTCTCTTGGAGCCAGCCCT 2487  
Db 648 AGCAAGGAGGAGTGTGGATGTGGAAACCTTACCTGTCTCTCTTGGAGCCAGCCCT 589  
Qy 2488 GCCAGAAACCCAGCTCACCAGAGCGGAAAGCAACCGTCCCGAGGAGGAACTCGAGCAG 2547  
Db 588 GCCAGAAACCCAGCTCACCAGAGCGGAAAGCAACCGTCCCGAGGAGGAACTCGAGCAG 529

QY	2548	CTGGAATAGAAATTAATTCCTCAACAGCCTGTCCCGAGCCATTTTCTCTGGAGGACGAGGAG	2607
Db	528	CTGGAATAGAAATTAATTCCTCAACAGCCTGTCCCGAGCCATTTTCTCTGGAGGACGAGGAG	469
QY	2608	CAAAATTCCTCGTGCCTCAGCATCGACAGCCTCTCCCTGTGGATGACAGTGAGAAGAAC	2667
Db	468	CAAAATTCCTCGTGCCTCAGCATCGACAGCCTCTCCCTGTGGATGACAGTGAGAAGAAC	409
QY	2668	CCATCAAAAGGCCTCTCAAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAGC	2727
Db	408	CCATCAAAAGGCCTCTCAAAGCTCGCGGACACCCCTGAGCTCAGGCGTACACTCTCTGGAGC	349
QY	2728	AGCCAGGCCGAGGCTCGAAGCTCCAGCTCGGAACATGGTGTGGCCCGGGGGGGCCCCACC	2787
Db	348	AGCCAGGCCGAGGCTCGAAGCTCCAGCTCGGAACATGGTGTGGCCCGGGGGGGCCCCACC	289
QY	2788	GACACCCCAAGCTATTTCATATGCTGTGAAAGTCCAAATACAGTCTCTTTAATGGTGAAACAC	2847
Db	288	GACACCCCAAGCTATTTCATATGCTGTGAAAGTCCAAATACAGTCTCTTTAATGGTGAAACAC	229
QY	2848	CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGGCATCAGCAGC	2907
Db	228	CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGGCATCAGCAGC	169
QY	2908	CAGATCCCAAGCTGCAGCCTTCAGCTTGGTCACCAAGAGCGGCGCCTGTTGCTACGAC	2967
Db	168	CAGATCCCAAGCTGCAGCCTTCAGCTTGGTCACCAAGAGCGGCGCCTGTTGCTACGAC	109
QY	2968	ATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTC	3027
Db	108	ATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTC	49
QY	3028	GCCTGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTAA	3075
Db	48	GCCTGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTAA	1

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:20:34 ; Search time 9684.08 Seconds  
(without alignments)  
18065.048 Million cell updates/sec

Title: US-09-155-676B-6  
Perfect score: 4596  
Sequence: 1 agcgggggactgtgccgtg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsl1:\*

9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3309.8	72.0	3971	3	CR749592 Homo sapi
2	2588.4	56.3	2844	9	AY413577 Homo sapi
3	2099.4	45.7	4174	3	AK038200 Mus muscu
4	2095.6	45.6	2765	9	AY413578 Pan trogl
5	1942.2	42.3	2829	9	AY413579 Mus muscu
6	1521.2	33.1	1640	3	CR612342 full-leng
7	943	20.5	1095	1	AL578455 AL578455
8	916	19.9	1133	1	AL556130 AL556130
9	876.2	19.1	1140	4	BM458601 AGENCOURT
10	811.4	17.7	955	5	BU541005 AGENCOURT
11	808.8	17.6	1038	4	BM454415 AGENCOURT
12	787.4	17.1	909	5	BX331085 BX331085
13	756.2	16.5	1037	4	BG775354 AGENCOURT
14	745	16.2	774	4	BG822015 AGENCOURT
15	741.6	16.1	2964	3	AK038828 Mus muscu
16	737.8	16.1	1166	4	BM458535 AGENCOURT
17	714.6	15.5	911	5	BU195099 AGENCOURT
18	705	15.3	762	4	B1259434 AGENCOURT
19	684.2	14.9	743	7	CA409541 AGENCOURT
20	675	14.7	702	5	EX102204 BX102204
21	665	14.5	741	4	BG685061 AGENCOURT
22	658.4	14.3	819	5	BQ231411 AGENCOURT
23	653.8	14.2	896	5	BU542238 AGENCOURT
24	647.8	14.1	707	1	AL042398 DKFZp434B

c	25	635.6	13.8	665	6	CB305346	UI-CF-EN1
	26	629.6	13.7	964	5	BQ923823	AGENCOURT
	27	622.2	13.5	991	4	BM550076	AGENCOURT
	28	605.8	13.2	702	4	BI489907	603031660
	29	604.4	13.2	768	4	BM048296	603625505
	30	596.2	13.0	806	4	BI161002	602865308
	31	591	12.9	701	4	BG774889	602849808
	32	580	12.6	702	7	CN409540	170006001
	33	568.6	12.4	702	7	CK979910	4111556 B
	34	564.2	12.3	580	5	BP267282	BP267282
	35	556.2	12.1	579	2	BE407122	601301805
	36	549.8	12.0	1172	4	BM543747	AGENCOURT
	37	548	11.9	556	2	BE276256	601144348
	38	538.4	11.7	795	4	BI826678	603077440
	39	530	11.5	736	2	BB632073	BB632073
	40	524.4	11.4	548	2	BE391253	601285947
	41	520	11.3	531	6	CA748591	UI-H-F11
	42	518.8	11.3	659	2	AM178916	MRO-ST001
	43	512.6	11.2	556	2	BE388034	601283733
	44	512.2	11.1	702	4	BI554785	603236762
	45	512.2	11.1	730	1	AI956650	UL79H03.Y

ALIGNMENTS

RESULT 1	CR749592	3971 bp	mrna	linear	HTC 19-AUG-2004
LOCUS	Homo sapiens mRNA; cDNA DKFZp686J04131 (from clone DKFZp686J04131).				
DEFINITION	CR749592				
ACCESSION	CR749592.1	GI:51476844			
VERSION	HTC				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3971)				
AUTHORS	Bloeker, H., Boecker, M., Brandt, P., Meves, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.				
CONSRM	The German cDNA Consortium				
TITLE	Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764				
JOURNAL	Neuherberg, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.				

This clone (DKFZp686J04131) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp686J04131>  
Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

FEATURES  
source  
1..3971  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686J04131"  
/tissue\_type="uterus"  
/clone\_lib="686 (synonym: hlcc3). Vector pSport1\_Sfi; host DH10B; sites SfiIA + SfiIB"  
/note="MAP3K14 protein, N-terminus truncated, differentially spliced"  
1..3971  
/gene="DKFZp686J04131"  
/gene="DKFZp686J04131"  
/codon\_start=1  
/product="hypothetical protein"

/protein\_id="CAH18391.1"  
/db\_xref="GI:51476845"  
/translation="CFKOLGEGRLPALPRSELHKLISPLQCLINHVKLHHPDGGPLP  
LPTHPPPHLPQPKPPLPESFLGKLACVDQKPLPDLHLSKLACVSPKPLPGRHLE  
PSCUSRGAHERFSEYELVHALQGSVSSGQASHLSLAKTWAARSGRSPSPKTEDN  
EGLVLETKLPVDYREYBEVHWATHQLRGRSGFGEVRMEDKQTFQCAVKKRVLEV  
FRAELMACAGLTSPRIVPLVCAVREGPWNI FMBLEGGSLGQLVKEQGLPEDRAL  
YVLAGALRGLEYLHSRRLLHGDVKADNVLLSDGSHAALCDFGHAVCIQPDGLGKSLI  
TGDXTPGNETHMAPVILGRSCDAKVDVWSSCCMMLHMLNGCHPWQTFRRGPLCIKVG  
DGTQHGGQBELQGPQAQAWRGCSAVADGARGGLVTHVGS"

ORIGIN

Query Match 72.0%; Score 3309.8; DB 3; Length 3971;  
Best Local Similarity 93.0%; Pred. No. 0;  
Matches 3673; Conservative 0; Mismatches 32; Indels 244; Gaps 10;  
QY 856 TGTTTAAAGCAGCTTGGGAGGGGCTACGGCCGGCTCTGCCTCGATCAGAACTCCACAAA 915  
DB 1 TGTTTAAAGCAGCTTGGGAGGGGCTACGGCCGGCTCTGCCTCGATCAGAACTCCACAAA 60  
QY 916 CTGATCAGCCCTTCAATGCTGAACACACGTTGTGAAACTGCACACCCCGCAGGACGGA 975  
DB 61 CTGATCAGCCCTTCAATGCTGAACACACGTTGTGAAACTGCACACCCCGCAGGACGGA 120  
QY 976 GGGCCCTGCGCTGCCACGACGCCCTTCCCTATAGCAGACTGCCTCATCCCTTCCCA 1035  
DB 121 GGGCCCTGCGCTGCCACGAC-----CCCTTCCCA 153  
QY 1036 TTCCACCTCTCAGCCCTTGAAAACCTCACCTCTGGAGTCTTCTGGGCAAACTGGCC 1095  
DB 154 TTCCACCTCTCAGCCCTTGAAAACCTCACCTCTGGAGTCTTCTGGGCAAACTGGCC 213  
QY 1096 TGTGTAGACAGCAGAAACCTTGGCTGACCCACACCTGAGCAAACTGSCCTGTGTAGAC 1155  
DB 214 TGTGTAGACAGCAGAAACCTTGGCTGACCCACACCTGAGCAAACTGSCCTGTGTAGAC 273  
QY 1156 AGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCTGTGGTGGCCAT 1215  
DB 274 AGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGTCTGTGGTGGCCAT 333  
QY 1216 GAGAGTTTCTGTGGAGAAATACCTAGTGCAATGCTCTGCAAGGAGCGTGAGCTCAAGC 1275  
DB 334 GAGAGTTTCTGTGGAGAAATACCTAGTGCAATGCTCTGCAAGGAGCGTGAGCTCAGGC 393  
QY 1276 CAGGCCACAGCTCAGCAGCTGSCCAAGACTCGGCAGCAGCGGGCTCCAGATCCCGG 1335  
DB 394 CAGGCCACAGCTCAGCAGCTGSCCAAGACTCGGCAGCAGCGGGCTCCAGATCCCGG 453  
QY 1336 GAGCCAGCCCCAAACTGAGGACAAACGAGGGTGTCTGCTCACTGAGAAACTCAAGCCA 1395  
DB 454 GGGCCACAGCCCCAAACTGAGGACAAACGAGGGTGTCTGCTCACTGAGAAACTCAAGCCA 513  
QY 1396 GTGGATTATGATACGAGAGAAAGTCCACTGGGCCACGCACAGCTCGCCTGGGCAGA 1455  
DB 514 GTGGATTATGATACGAGAGAAAGTCCACTGGGCCACGCACAGCTCGCCTGGGCAGA 573  
QY 1456 GGCTCTCTCGGAGAGTGACAGGATGGAGGACAGCAGCTGCTCCAGTCCGCTGTC 1515  
DB 574 GGCTCTCTCGGAGAGTGACAGGATGGAGGACAGCAGCTGCTCCAGTCCGCTGTC 633  
QY 1516 AAAAAGGTGCGCTGGAAGTATTTCCGGGACAGGAGCTGATGGCATGTGCAGGATTGACC 1575  
DB 634 AAAAAGGTGCGCTGGAAGTATTTCCGGGACAGGAGCTGATGGCATGTGCAGGATTGACC 693  
QY 1576 TCACCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGCCCTTGGGTCAACATCTTC 1635  
DB 694 TCACCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGCCCTTGGGTCAACATCTTC 753  
QY 1636 ATGGAGCTGCTCGAAGGTGGCTCCCTGGCCAGCTGCTCAAGGACAGGGCTGTCTCCCA 1695  
DB 754 ATGGAGCTGCTCGAAGGTGGCTCCCTGGCCAGCTGCTCAAGGACAGGGCTGTCTCCCA 813  
QY 1696 GAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTGGAAATACCTCACTCA 1755

DB 814 GAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTTGGAGGGTCTGGAAATACCTCCACTCA 873  
QY 1756 CGAAGGATTTCTGCATGGGACGTCAAAGCTCACAACGTCCTCTGTCTCAGCGATGGGAGC 1815  
DB 874 CGAAGGATTTCTGCATGGGACGTCAAAGCTCACAACGTCCTCTGTCTCAGCGATGGGAGC 933  
QY 1816 CACGAGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCTCAACCTGTATGGCTTGGAAAG 1875  
DB 934 CACGAGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCTCAACCTGTATGGCTTGGAAAG 993  
QY 1876 TCCTTGTCTACAGGGGACTACATCCCTGGGACAGAGACCCACATGGCTCCGAGAGTGTG 1935  
DB 994 TCCTTGTCTACAGGGGACTACATCCCTGGGACAGAGACCCACATGGCTCCGAGAGTGTG 1053  
QY 1936 CTGGGAGGAGCTCGAGCCCAAGTGATCTCTGGAGCAGCTGCTGTATGATCTGCAC 1995  
DB 1054 CTGGGAGGAGCTCGAGCCCAAGTGATGCTCTGGAGCAGCTGCTGTATGATCTGCAC 1113  
QY 1996 ATGCTCAACGGCTGCCACCCCTTGGACTCAGTTCTTCCGAGGGCCGCTCTGCCTCA 2050  
DB 1114 ATGCTCAACGGCTGCCACCCCTTGGACTCAGTTCTTCCGAGGGCCGCTCTGCCTCAAGT 1173  
QY 2051 ----- 2050  
DB 1174 GGTGACGGCACCCAGCATGGCCAAACAGGGGGAGCTGCAGGGCCCTGGGGCCAGGCTGCA 1233  
QY 2051 ----- 2050  
DB 1234 GGCTGGAGGGCTGCTCTGCGGTGGCAGATGGAGCAAGGGAGGCTTGGTCAATCCATGTG 1293  
QY 2051 ----- 2050  
DB 1294 GGATCCTTAATTCAAATACTTCCACCCCAAATCTGGGGGTACAAGGGGAGATCCAGCTGG 1353  
QY 2051 -----AGATTGCCAGCGAGCTTCCGCTGTGTAGGGAGATCCAC 2089  
DB 1354 GAGGCTCACCAAGTTTCATGGCAGATTGTCAGCGAGCTTCCGCTGTGTAGGGAGATCCAC 1413  
QY 2090 CCTCTGCGCCCTCTCACAGCCAGGCCATCCTAAGAGGGCTGAGGAAAGAGCCCATCC 2149  
DB 1414 CCTCTGCGCCCTCTCACAGCCAGGCCATCCTAAGAGGGCTGAGGAAAGAGCCCATCC 1473  
QY 2150 ACCGCTGTCTGCAGCGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAG 2209  
DB 1474 ACCGCTGTCTGCAGCGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAG 1533  
QY 2210 GTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAAATCAAG 2269  
DB 1534 GTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACGCCAAATCAAG 1593  
QY 2270 CCNATTACCAAGACCCCTCCATGCCAGCGAGAGCTTTTCGCCAAGGGCCCGCAGGCG 2329  
DB 1594 CCAATTACCAAGACCCCTCCATGCCAGCGAGAGCTTTTCGCCAAGGGCCCGCAGGCG 1653  
QY 2330 CCGGCCAGCTGAGGAGACAAACAGCAGAGCCCTTAAGCTCCAGGCTCTCTCCACCCAG 2389  
DB 1654 CCGGCCAGCTGAGGAGACAAACAGCAGAGCCCTTAAGCTCCAGGCTCTCTCCACCCAG 1713  
QY 2390 AGCCCCAGAGCCAAACAAAGTCTCTCTCTTGAATTTTGAAGAGGAGTCTGGGATGT 2449  
DB 1714 AGCCCCAGAGCCAAACAAAGTCTCTCTCTTGAATTTTGAAGAGGAGTCTGGGATGT 1773  
QY 2450 GGGAAACCTTACCTCTGTCTCTCTGGAGCCAGCCCTTCCGCAAGAACCCAGCTCACCAG 2509  
DB 1774 GGGAAACCTTACCTCTGTCTCTCTGGAGCCAAACCCCTTCCGCAAGAACCCAGCTCACCAG 1833  
QY 2510 AGCGAAAGAACCCCTTCCCGGAGCAGGAACTGCGAGAGTGGAAATAGAAATATTCTCTCA 2569  
DB 1834 AGCGAAAGAACCCCTTCCCGGAGCAGGAACTGCGAGAGTGGAAATAGAAATATTCTCTCA 1893  
QY 2570 ACAGCTGTCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTGTCGCTCAGCA 2629

1894	ACAGCCTGTC	CCAGCCATTTT	TCTCTGGAGGAGCAGGAACAAA	TTTCTCTCTGTCGCTCAACA	1955
2630	TCGACAGCCT	TCCCTCTG	CGGATGACAGT	GAGAGAAACCCATCAAAAGCCCTCTCAAAAGCT	2689
1954	TCGACAGCCT	TCCCTCTG	TCGGATGACAGT	GAGAGAAACCCATCAAAAGCCTCTCAAAAGCT	2013
2690	CGCGGGACA	CCCTGAGCT	CAGGCGTACACTCT	GTGGAGAGCAGGCCCGAGGCTCGAAAGCT	2749
2014	CGCGGGACA	CCCTGAGCT	CAGGCGTACACTCT	GTGGAGAGCAGGCCCGAGGCTCGAAAGC	2073
2750	CCAGCTGGAA	CATGCTGTG	TGCGCGGGGGCGCCACCG	CACACCCCAAGCTATTTCGAATG	2809
2074	CCAGCTGGAA	CATGCTGTG	TGCGCGGGGGCGCCACCG	CACACCCCAAGCTATTTCGAATG	2133
2810	GTGTGAAGT	CCAAATACAGT	CTCTTAATGGTGAAACACTG	CGCATCCCGGGAGTTCACCC	2869
2134	GTGTGAAGT	CCAAATACAGT	CTCTTAATGGTGAAACACTG	CGCATCCCGGGAGTTCACCC	2193
2870	GGGTCAAAGT	GGGAGACATG	CGCCTATCGCAGC	CGCAGATCCCAAGCTTCGAGCCTTCA	2929
2194	GGGTCAAAGT	GGGAGACATG	CGCCTATCGCAGC	CGCAGATCCCAAGCTTCGAGCCTTCA	2253
2930	GCTTGTGT	CAACAAAGAG	CGGCGAGCCTGTTCGT	TACGACATGGAGGTGCCAGACTCGGGCA	2989
2254	GCTTGTGT	CAACAAAGAG	CGGCGAGCCTGTTCGT	TACGACATGGAGGTGCCAGACTCGGGCA	2313
2990	TCGACCTG	CGAGTGACACTG	CGCCCTGATGGCAGCT	TTCGGCTGGAGCTGGAGGCTCAAGC	3049
2314	TCGACCTG	CGAGTGACACTG	CGCCCTGATGGCAGCT	TTCGGCTGGAGCTGGAGGCTCAAGC	2372
3050	ATGGCCAG	CTGGAGAA	CAGGCCCTTAACCTTG	CCCTCCACCGCCGGCTCCACACTGCCGGA	3109
2373	ATGGCCAG	CTGGAGAA	CAGGCCCTTAACCTTG	CCCTCCACCGCCGGCTCCACACTGCCGCG-	2431
3110	AAGCAGCCT	TCTCTGCT	CGGTGCAGATGCTCCCTG	TAACAAACAGGCTCAGCGGTCCCA	3169
2432	AAGCAGCCT	TCTCTGCT	CGGTGCAGATGCTCCCTG	-AAACACAGGCTCAGCGGTCCCA	2490
3170	GGGGATTG	CCAGCCCCCGGCT	CACAGTGGGAAACAGGGCCCT	CGCAGCAGCAAGAGTGGGG	3229
2491	GGGATCTG	CCAG--	CCCCCGGCTCAGCAGTGGGAC	CCAGGGCTCGCAGCAGCAGAGTGGGG	2549
3230	GCAAGCAGAA	TGCTCC	CAGATTTCACACTG	AGCCCTGCCCACTCTGTAATAAAC	3289
2550	GCAAGCAGAA	TGCTCC	CAGATTTCACACTG	AGCCCTGCCCACTCTGTAATAAAC	2609
3290	A-TCGCGCA	CGTGAAGAGACAG	-AGSAGGATGGCAGGAGTTAC	CTCGGGGAAACAAACAG	3348
2610	ACTCGCGCA	CGTGAAGAGACAG	-AGSAGGATGGCAGGAGTTAC	CTCGGGGAAACAAACAGG	2668
3349	GGATCTTTT	TCTGCCCCCTGCT	TCCAGTTCAGTTGGCCCTG	ACCCGCTTCGGATCAGTGACCAT	3408
2669	----	ATCTTCTCTCCCTGCT	TCCAGTTCAGTTGGCCCTG	ACCCGCTTCGGATCAGTGACCAT	2724
3409	TTGTTGGC	AGACAGGGAGACAGCT	TTCCAGCCTGGGTTCAGAA	GGGGTGGCGAGCCCTT	3468
2725	TTGTTGGC	AGACAGGGAGACAGCT	TTCCAGCCTGGGTTCAGAA	GGGGTGGCGAGCCCTT	2784
3469	CGGCCCCCT	CACCTTCAGGCTGCT	GTGAGAGTGTCAAGTGTGTA	AGGGGCCCAAACTCAGG	3528
2785	CGGCCCCCT	CACCTTCAGGCTGCT	GTGAGAGTGTCAAGTGTGTA	AGGGGCCCAAACTCAGG	2844
3529	TTTCAGTGC	AGAACACAGGTTCAG	CAGGATATGCCCGCCGCT	TAGGTTAAGGGGGCCCTCTAAAC	3588
2845	TTTCAGTGC	AGAACACAGGTTCAG	CAGGATATGCCCGCCGCT	TAGGTTAAGGGGGCCCTCTAAAC	2904
3589	CCCTTGCT	TGCTTCACTCGGCCAGCT	CAACCCCTTTTGGGTGTAG	GGGAAAGAAATGCCT	3648
2905	CCCTTGCT	TGCTTCACTCGGCCAGCT	CAACCCCTTTTGGGTGTAG	GGGAAAGAAATGCCT	2964
3649	GACCTGGG	AAGGCTCCCTGGT	TAGATACACCACTTTT	TTCAGGTTGTGCAACACAGGT	3708
2965	GACCTGGG	AAGGCTCCCTGGT	TAGATACACCACTTTT	TTCAGGTTGTGCAACACAGGT	3024

QY	3709	CCTGAGTTGACCTCTGTTT	CAGCCAAAGGACCAAGAAAGAGTGTGTAACTGAAGTGGTTCTC	3768
DB	3025	CCTGAGTTGACCTCTGTTT	CAGCCAAAGGACCAAGAAAGAGTGTGTAACTGAAGTGGTTCTC	3084
QY	3769	AGTCCCAGACATGTGCCCTTT	GCTGCTGGCTACCACTCTTCCCAGAGCAGCAGGCC	3828
DB	3085	AGTCCCAGACATGTGCCCTTT	GCTGCTGGCTACCACTCTTCCCAGAGCAGCAGGCC	3144
QY	3829	CGAGCCCTTCAGGCC	CAGCACTGCCCGACACTGCTGGGCACTCAGTTCCCTCATCTGTGA	3888
DB	3145	CGAGCCCTTCAGGCC	CAGCACTGCCCGACACTGCTGGGCACTCAGTTCCCTCATCTGTGA	3204
QY	3889	AAGTGAAGGGTGATGCAGGATAT	GCCTGTGACAGAAACAGTCTGTGGATGACATGATCAG	3948
DB	3205	AAGTGAAGGGTGATGCAGGATAT	GCCTGTGACAGAAACAGTCTGTGGATGACATGATCAG	3264
QY	3949	TGCTAAGGAAGCAGCAGAGAGAGACG	-TCCGGCCGCCAGCCCTACTATCAGTGTCCAG	4007
DB	3265	TGCTAAGGAAGCAGCAGAGAGAGACG	-TCCGGCCGCCAGCCCTACTATCAGTGTCCAG	3324
QY	4008	CGTCTCGTTCCCCAGAGCA	CAGCTCAGCATCACACTTGACACTCACCCCTGCCCTGCCCT	4067
DB	3325	CGTCTCGTTCCCCAGAGCA	CAGCTCAGCATCACACTTGACACTCACCCCTGCCCTGCCCT	3384
QY	4068	GGCCAGAGGGTACTTGC	CGCAGCGGCACTTGTGACTCTGATGACCTCAAAAGCACTTTTCATGGC	4127
DB	3385	GGCCAGAGGGTACTTGC	CGCAGCGGCACTTGTGACTCTGATGACCTCAAAAGCACTTTTCATGGC	3444
QY	4128	TGCCCTCTGGCAGGGCAGGGC	AGTGACACTGTAGGAGCATGCAAGCCAGAGAGAT	4187
DB	3445	TGCCCTCTGGCAGGGCAGGGC	AGTGACACTGTAGGAGCATGCAAGCCAGAGAGAT	3504
QY	4188	GGGGTGAAGGACACAGTCTT	TGAGCTGTGCACATGCACTGTGACTCTCAAACTCTTCCA	4247
DB	3505	GGGGTGAAGGACACAGTCTT	TGAGCTGTGCACATGCACTGTGACTCTCAAACTCTTCCA	3564
QY	4248	GATTTCTCTAAGAATAGCAC	CCCCCTTCCCATTGCCCCAGCTTATAGCCCTTTCTCCCAGGG	4307
DB	3565	GATTTCTCTAAGAATAGCAC	CCCCCTTCCCATTGCCCCAGCTTATAGCCCTTTCTCCCAGGG	3624
QY	4308	GAGCTACTCAGGACTCAGT	TAGCATTAATTCAGCTGTGAATCTGTCAGGGGGTGTCTGCTA	4367
DB	3625	GAGCTACTCAGGACTCAGT	TAGCATTAATTCAGCTGTGAATCTGTCAGGGGGTGTCTGCTA	3684
QY	4368	GCCTCAACTCCTGGGGCAGGGAC	GGCAGAGACTCCGTGGGAGAGCTCAATCCCAATC	4427
DB	3685	GCCTCAACTCCTGGGGCAGGGAC	GGCAGAGACTCCGTGGGAGAGCTCAATCCCAATC	3744
QY	4428	TTGCAAGACAGCCTTTGT	CTCAGCTGTCCAATTGAGTCTCAGCTGCTCCCGGGAGAGAG	4487
DB	3745	TTGCAAGACAGCCTTTGT	CTCAGCTGTCCAATTGAGTCTCAGCTGCTCCCGGGAGAGAG	3804
QY	4488	CCCCGGCCCCCAGCACATAAG	AACTCAGCCTTGGTACTGACAGTCTGGGTTGTAGAG	4547
DB	3805	CCCCGGCCCCCAGCACATAAG	AACTCAGCCTTGGTACTGACAGTCTGGGTTGTAGAG	3864
QY	4548	AACTCTTTTGAAGCAATAA	AGTTGGGGTGATGACAAATGTTAAAAAA	4596
DB	3865	AACTCTTTTGAAGCAATAA	AGTTGGGGTGATGACAAATGTTAAAAAA	3913

RESULT 2	AY413577	2844 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	AY413577				
DEFINITION	Homo sapiens MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,				
	AY413577				
ACCESSION	Homo sapiens genomic survey sequence.				
VERSION	AY413577				
KEYWORDS	AY413577.1	GI:39769539			
SOURCE	GSS.				
	Homo sapiens	(human)			
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 2844)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Perriera,S., Wang,G., Zheng,X.H., White,F.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
2 (bases 1 to 2844)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
Location/Qualifiers  
source  
1..2844  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
gene  
<1..>2844  
/gene="WAP3K14"  
/locus\_tag="HCM4937"  
ORIGIN  
Query Match 56.3%; Score 2588.4; DB 9; Length 2844;  
Best Local Similarity 91.1%; Pred. No. 0;  
Matches 2592; Conservative 0; Mismatches 252; Indels 0; Gaps 0;  
QY 232 ATGGCAGTGATGGAATGCGCTGCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAG 291  
DB 1 ATGGCAGTGATGGAATGCGCTGCCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAG 60  
QY 292 GAATCCCAAGCCAAAGAGAGAGCGCCGACCTGGGGGAGAAACAGAGCTCCCTCTAC 351  
DB 61 GAATCCCAAGCCAAAGAGAGAGCGCCGACCTGGGGGAGAAACAGAGCTCCCTCTAC 120  
QY 352 AAGCTTGAGCGCGTGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAAATGAC 411  
DB 121 AAGCTTGAGCGCGTGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAAATGAC 180  
QY 412 GTGATTACCAAGGGCAGACCCAGGAAGGCTCCGAGGCGAGGCCAGCTGCCATCTATTC 471  
DB 181 GTGATTACCAAGGGCAGACCCAGGAAGGCTCCGAGGCGAGGCCAGCTGCCATCTATTC 240  
QY 472 ATCGCCCGAGCTGAGTGTGAGAAATAGCCAAAGATTTCAGCCCAACCTTTTCAGAACGCAAT 531  
DB 241 ATCGCCCGAGCTGAGTGTGAGAAATAGCCAAAGATTTCAGCCCAACCTTTTCAGAACGCAAT 300  
QY 532 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACCAAT 591  
DB 301 TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACCAAT 360  
QY 592 GTGGGCCATGTCTACAGAGGCGAAAATGGCCCTGTGTGTTGGAAGGGAAAGCGTCGCAGC 651  
DB 361 GTGGGCCATGTCTACAGAGGCGAAAATGGCCCTGTGTGTTGGAAGGGAAAGCGTCGCAGC 420  
QY 652 AAAGCCCGGAAGAAACGGAAGAGAGAGCTCAAAAGTCCCTGGCTCATGCGAGGTGGCC 711  
DB 421 AAAGCCCGGAAGAAACGGAAGAGAGAGCTCAAAAGTCCCTGGCTCATGCGAGGTGGCC 480  
QY 712 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGCTGCCACCATCCCACTGCGAGGAG 771  
DB 481 TTGGCCAAACCCCTCCCAAGGACCCCTGAGCAGGAGAGCTGCCACCATCCCACTGCGAGGAG 540  
QY 772 GATGAGTCTCCACTCGGCGCCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 831  
DB 541 GATGAGTCTCCACTCGGCGCCCCATATGTTAGAAACACCCCGAGTTCACCAAGCCTCTG 600

QY 832 AAGGAACAGGCGCTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCTACGGCGGGCT 891  
DB 601 AAGGAACAGGCGCTTGGGNN 660  
QY 892 CTGCGCTCGATCAGAACTCCAACAACTGATCAGCCCTTGGCAATGTCTGAACACACCTGTGG 951  
DB 661 NNN 720  
QY 952 AAATGTCACCAACCCCGAGGACGGAGGCCCTTGGCCCTTGGCCACGACGCCCTTCCCTAT 1011  
DB 721 AAATGTCACCAACCCCGAGGACGGAGGCCCTTGGCCCTTGGCCACGACGCCCTTCCCTAT 780  
QY 1012 AGCAGACTGCTCATCCCTTCCATTCCACCTCTCCAGCCCTGGAACCTCACCCTCTG 1071  
DB 781 AGCAGACTGCTCATCCCTTCCATTCCACCTCTCCAGCCCTGGAACCTCACCCTCTG 840  
QY 1072 GAGTCTTTCTGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTCCCTGACCCACAC 1131  
DB 841 GAGTCTTTCTGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTCCCTGACCCACAC 900  
QY 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTTGGCCGACACCTGAGGCC 1191  
DB 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTTGGCCGACACCTGAGGCC 960  
QY 1192 AGTGCCTGTCTCGTGGTCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1251  
DB 961 AGTGCCTGTCTCGTGGTCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCT 1020  
QY 1252 CTGCAAGCAGCGTGAGCTCAAGCCGACGCCCAAGCTGACAGCTGGCCAAAGACCTGG 1311  
DB 1021 CTGCAAGCAGCGTGAGCTCAAGCCGACGCCCAAGCTGACAGCTGGCCAAAGACCTGG 1080  
QY 1312 GCAGCACGGGCTCCAGATCCGGGAGCCGAGCCCAAACTGAGAGCAACGAGGGTCTC 1371  
DB 1081 GCAGCAAGGGGCTCCAGATCCGGGAGCCGAGCCCAAACTGAGAGCAACGAGGGTCTC 1140  
QY 1372 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATTACCGAGAAAGTCCACCTGGGCC 1431  
DB 1141 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATTACCGAGAAAGTCCACCTGGGCC 1200  
QY 1432 AGCCACAGCTCCGCTGGGCGAGGCTCTTTCGGAGAGGTGCACAGATGGAGACAAAG 1491  
DB 1201 AGCCACAGCTCCGCTGGGCGAGGCTCTTTCGGAGAGGTGCACAGATGGAGACAAAG 1260  
QY 1492 CAGACTGCTTCCAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGCGAGGAG 1551  
DB 1261 CAGACTGCTTCCAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTTCGGGCGAGGAG 1320  
QY 1552 CTGATGGCATGTGACAGGATTGACCTCACCCAGAAATTTGCCCTTGTATGGAGCTGTGAGA 1611  
DB 1321 CTGATGGCATGTGACAGGATTGACCTCACCCAGAAATTTGCCCTTGTATGGAGCTGTGAGA 1380  
QY 1612 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671  
DB 1381 GAAGGCGCTTGGGTCAACATCTTCATGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440  
QY 1672 GTCAAGGACGAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1731  
DB 1441 GTCAAGGACGAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1500  
QY 1732 GAGGCTCTGGAATACCTCCACTCAGAAAGGATTCTGCATGGGAGAGCTCAAAAGCTCACAAC 1791  
DB 1501 GAGGCTCTGGAATACCTCCACTCAGAAAGGATTCTGCATGGGAGAGCTCAAAAGCTCACAAC 1560  
QY 1792 GTGCTCTGTCCAGCGATGGGAGCCACCGAGCCCTCTGTGACTTTGGCCCATGCTGTGTGT 1851  
DB 1561 GTGCTCTGTCCAGCGATGGGAGCCACCGAGCCCTCTGTGACTTTGGCCCATGCTGTGTGT 1620  
QY 1852 CTTCAACTGATGGCTGGAAAGTCTTGTGTACAGGGGACTACATCCCTGGGCACAGAG 1911  
DB 1621 CTTCAACTGATGGCTGGAAAGTCTTGTGTACAGGGGACTACATCCCTGGGCACAGAG 1680

QY 1912 ACCCACATGGCTCCGGAGGTGGTCTGGGAGGAGCTGCGACGCCAAGGTGGATGTCTGG 1971  
Db 1681 ACCCACATGGCTCCGGAGGTGGTCTGGGAGGAGCTGCGACGCCAAGGTGGATGTCTGG 1740  
QY 1972 AGCAGCTCTGTATGATGCTGCATGCTCAACGCTGCGACGCTGCGACCTGCGACCTCTTC 2031  
Db 1741 AGCAGCTCTGTATGATGCTGCATGCTCAACGCTGCGACGCTGCGACCTCTTC 1800  
QY 2032 CGAGGGCGCTCTGCCTCAAGATTGCCAGCGAGCTTCGCGCTGTGAGGGAGATCCCAACC 2091  
Db 1801 CGAGGGCGCTCTGCCTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1860  
QY 2092 TCCTGCGCCCTCTCACAGCCAGCCATCAAGAGGGCTGAGAAAGACCCATCCAC 2151  
Db 1861 NNN 1920  
QY 2152 CGCGTGTCTGACGCGAGCTGGAGGGAAGTGAACCGGCACTACAGCAAGTGGAGGT 2211  
Db 1921 NNN 1980  
QY 2212 CTGAAGAGCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2271  
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2040  
QY 2272 AATTACCAACAGACCTCCATGCCAGCCGAGAGAGCTTTGCCAAGGGCCCCAGGGCCC 2331  
Db 2041 AATTACCAACAGACCTCCATGCCAGCCGAGAGAGCTTTGCCAAGGGCCCCAGGGCCC 2100  
QY 2332 CGGCCAGTGGAGAGACAAAGGAGGAGCCCTAAGCTCCAGCTCTCTCCACCCAGAG 2391  
Db 2101 CGGCCAGTGGAGAGACAAAGGAGGAGCCCTAAGCTCCAGCTCTCTCCACCCAGAG 2160  
QY 2392 CCCCCAGAGCAAAAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2451  
Db 2161 CCCCCAGAGCAAAAGTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGG 2220  
QY 2452 GAACCTTACTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCCAGAG 2511  
Db 2221 GAACCTTACTCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCCAGAG 2280  
QY 2512 CGGAAGCAACCGTCCGGAGCAGACTGAGCAGCTGGAATAGATTTCTCTCAAC 2571  
Db 2281 CGGAAGCAACCGTCCGGAGCAGACTGAGCAGCTGGAATAGATTTCTCTCAAC 2340  
QY 2572 AGCTGTCTCCAGCAATTTCTCTGAGGAGCAGGACAAATTTCTCTGCTCTCAGCATC 2631  
Db 2341 AGCTGTCTCCAGCAATTTCTCTGAGGAGCAGGACAAATTTCTCTGCTCTCAGCATC 2400  
QY 2632 GACAGCTCTCTCTCTGATGACAGTGAGAGAACCCATCAAGGCTCTCTCAAGCTCG 2691  
Db 2401 GACAGCTCTCTCTCTGATGACAGTGAGAGAACCCATCAAGGCTCTCTCAAGCTCG 2460  
QY 2692 CGGACACCTCTGAGCTCAGGCTACATCTCTGGAGCAGCGCCGAGGCTCGAAGCTCC 2751  
Db 2461 CGGACACCTCTGAGCTCAGGCTACATCTCTGGAGCAGCGCCGAGGCTCGAAGCTCC 2520  
QY 2752 AGCTGGAACATGTGTGCGCCCGGGGGCCGACCCAGACCCCAAGCTATTTCAATGGT 2811  
Db 2521 AGCTGGAACATGTGTGCGCCCGGGGGCCGACCCAGACCCCAAGCTATTTCAATGGT 2580  
QY 2812 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGAGGAGTCCACCGG 2871  
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGGTGAACACCTGCAATCCGAGGAGTCCACCGG 2640  
QY 2872 GTCAAGTGGAGACATCGCCATCGGCATCAGCAGCGAGATCCAGCTGCGAGCTTCAGC 2931  
Db 2641 GTCAAGTGGAGACATCGCCATCGGCATCAGCAGCGAGATCCAGCTGCGAGCTTCAGC 2700  
QY 2932 TTGGTTCACCAAGAGCGGCGCTGCTTCCGTACCAATGAGAGTCCAGACTCCGGGATC 2991  
Db 2701 TTGGTTCACCAAGAGCGGCGCTGCTTCCGTACCAATGAGAGTCCAGACTCCGGGATC 2760  
QY 2992 GACCTGAGTGCACATGGGCCCTCTGATGGCAGCTTTCGCTGGAGCTGGAGGCTCAAGCAT 3051

Db 2761 GACCTGAGTGCACATGGGCCCTGATGGCAGCTTCGCTGGAGCTGGAGGCTCAAGCAT 2820  
QY 3052 GGCCAGCTGGAGAACAGCCCTTAA 3075  
Db 2821 GGCCAGCTGGAGAACAGCCCTTAA 2844  
RESULT 3  
AK036200  
LOCUS  
DEFINITION  
Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:9630045621 product:mitogen-activated  
protein kinase kinase 14, full insert sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AK036200.1 GI:26331227  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
PUBMED  
10349636  
REFERENCE  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
PUBMED  
11042159  
REFERENCE  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
PUBMED  
11076861  
REFERENCE  
4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
REFERENCE  
5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
REFERENCE  
6  
(bases 1 to 4174)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohashi, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
TITLE

## JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

## FEATURES

Location/Qualifiers

source

1..4174  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:9630045G21"  
/db\_xref="taxon:10090"  
/clone="9630045G21"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="16 days neonate"  
85..2913  
/note="unnamed protein product; mitogen-activated protein kinase kinase kinase 14 (MGD|MGI:1858204, GB|NM\_016896, evidence: BLASTN, 100%, match=2829)  
putative"  
/codon\_start=1  
/protein\_id="BAC29144.1"  
/db\_xref="GI:26331228"  
/translation="MAWMEVACPPTPGSANGVQKELAKAKETQSLGKKQSCIFKLEA  
AGSQYVSSESIDQIPNNVAHATEGMARVCRGRHGRKARKRKRSKSLAQAGVA  
LAKPLRPTPEQSSCTIPQVEDSPLNLYARNVSQFTKPLGGPGLHLGKQKDEGLR  
LVPLRPELHLKLSPLQCLNHWKHLHPQATGPRPHTPFPYSGMHPFPFVPLEPWK  
PYMLDSALVDLKLAVSGORPLPCPHLSQLAHGDSQKPLPGPHLESSCPSSGALEKVP  
VVEYLVAHQSGVSGQSHSLASLAKTWSGSSAKLQRLGPETEDNEGVELTEKLKVPD  
YEYREVEHMTQHPVRKGSFGFVHRMDKQTFQCAVKKRVLEFRVEELVACAGLS  
SPRIPLFVAVRGEPWNAIFMELLEGSLGLQIKQMGCLPEDRALYILQALLEGLEYL  
HTRIRLHGDKADNVLSSDSCMMHLNGCHPWTOYFRGPLCLKIASBPPIREIPSC  
PVMWGPCKDAKVDIWSGCCMMHLNGCHPWTOYFRGPLCLKIASBPPIREIPSC  
APLTAQAQIEGLRKEPVHRASAMELRRXVKALQEVGGLKSPWKEVKEPRPPQDQA  
TCHQTLPTPPNPAPKANTDCAPEPQPLPEPPEPSKAPALNLSKESGTWEPLPL  
SLSDPATAKPSFPORRATLPELELQLELFLNLSLQPPSLEBEQILSLCSIDSL  
SLSDSEKNPSKASOSRDTLSSGVHWSNQAEARTCSCTALARGPTDIPSPYNGV  
KVQIQSLNGEHLHIREFHRVKVGDITGTISQIPATFSLVTKDQRPVCDWEVPSDG  
IDLOCTLAPDGSFAMTRVKGHQLNRP"  
4157..4162  
/note="putative"  
4174  
/note="putative"

## CDS

381 GTTCTGCGGAAAGTGGGAGATCCTCAATGACGTGATTACCAAGGCACAGCCAGGAAGG 440  
234 GTTCTGTGGGAAGTGGGAGATCCTAAACGACGTGATACCAAGGCACAGCCAGGAAGC 293  
441 TTCCGAGGACGAGGCGAGCTGCGCATCTCTATCATCGCCAGGCTGAGTGTGAGAATAGCCA 500  
294 CTCTGAGGAGGAGACACCGGCCATCTCCATCATCGCCAGGCTGAATGTGAGAATAGCCA 353  
501 AGAGTTTACGCCCCACCTTTTTCAGAACGATTTTTCATCGCTGGTGCACAAACAGTACAGCCA 560  
354 AGAGTTTACGCCCCACCTTTTTCAGAACGATTTTTCATCGCTGGTGCACAAACAGTACAGCCA 413  
561 GTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCTACAGAGGCGCAAAATGCG 620  
414 GTCTGAGAGTCTCGATCAAAATCCCCCAACATGTGGCCCATGCACTGAGGCGCAAAATGCG 473  
621 CCGTGTGTGTTGGAAAGGGAACGCTCGCAGCAAAAGCCCGGGAAGAAACAGGAAGAAGAG 680  
474 CCGTGTGTGCGGAGGGGAAAACGTCACGGCAAAAGCCCGGGAAGAAACAGTACGAAGAGAG 533  
681 CTCAAAGTCCCTGGCTCATGACAGGAGTGGCCCTTGGCGAAAACCCCTCCCGAGGACCCCTGA 740  
534 CTCGAAGTCACTGGGCCACGAGGAGTGGCCCTTAGCCAAAGCCCTTGGCCAGAACCCCTGA 593  
741 GCAGGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCCACCTCGCGGCCCATATGTT 800  
594 GCAAGAGAGTGTACCATCCAGTACAGGAAGATGAGTCTCCACTAGGCAACCTCTATGTC 653  
801 TAGAAACACCCCGCAGTCTCACAAGCCCTCTGAAAGGAACACGAGCCCTTGGGCAACCTCTGTTT 860  
654 CAGAAATGCTCCCAAGTTTCAACAAGCCCTCTGGGGGACACGAGCCCTTGGCCACCTGTGTT 713  
861 TAAGCAGCTTGGCGAGGCGCTACGGCCGGCTCTGGCTCGATCAGAACTCCCAAACTGAT 920  
714 TAAGAAACAAATGAAGCCCTTGGCAACCGGTACTGCTCGACAGAACTCCCAAACTGAT 773  
921 CAGCCCTTGCATATGCTGAAACCACTGTGAAACCTGCAACACCCAGGAGCGAGGCGCC 980  
774 CAGCCCTTGCATATGCTGAAACCACTGTGAAACCTGCAACACCCAGGAGCGCAAGGCGCC 833  
981 CCGTCCCTGCGCCAGGACACCCCTTCCCTATAGCAGACTGCTCATCTCCCTTCCCAATCCA 1040  
834 CCGGCCCCACCGCATCAACCCCTTCCCTACAGGGAATGCCCCATCTTTCCTTCCCAATCTA 893  
1041 CCCTCTCAGCCCTTGAAACCTCAACCTCTCGAGT---CCTTCTTGGGCAACCTGCGCTG 1097  
894 CCGCTTGGAGCCCTTGGAAACCTTATATGCTGAGACTTGTCCGCTCTTGGCAACCTAGCGCG 953  
1098 TGTAGACAGCCAGAAACCCCTTGCCTGACCCCA---CACCTGAGCAAACTGGGCTGTGTAGA 1154  
954 TGTAGCGGCCAGCGGCCCTCTGCTGGCCACCGCATCTAAGCCCACTGGGCCCATGGAGA 1013  
1155 CAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGCTGCGCTGTCTGTGTGTCGCA 1214  
1014 CAGTCAGAAAGCCCTGCTGCGCCACACCTGGAGTCCAGCTGCGCTCTCGGGGTGCGCT 1073  
1215 TGAGAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGCGACGCTGAGCTCAAG 1274  
1074 AGAAAAGTTTCCCGTGGAGGAATACCTGTGTGATGCGCTTCCAAAGGAAGTGTGAGCTCAGG 1133  
1275 CCAGGCCACAGCTGACAGCCCTGGCCCAAGACCTGGGCGAGCAGGGGCTCCAGATCCCG 1334  
1134 CCAGGCCACAGCTGCGCCAGCTGGCTTGAACATGGTCTCTCGGGAAGCGCCAGCTGCA 1193  
1335 GGAGGCCAGCCCCAAAACCTGAGGACACAGGGGTGTCTGTCTCATGTGAGAACTCAAGCC 1394  
1194 GAGGCTCGGCCCCGAAAACCTGAGGACACAGGGGGTCTCTGTCTTACTGAGAACTCAAGCC 1253  
1395 AGTGATTTATCAGTACCGAGGAAGTCCACTGTGGCCACCGACCGCTCCCGCTGGGGCAG 1454  
1254 AGTGATTTATGATATCGAAGAGAGTCCACTGTGATGACACACCGCTCGGGTGGGCGAG 1313  
1455 AGGCTCCTTTCGAGAGGTGCAAGGATGGAGGACAAAGCAGACTGGCTTCCAGTGCCTGT 1514

polyA\_signal

polyA\_site

## ORIGIN

Query Match 45.7%; Score 2099.4; DB 3; Length 4174;  
Best Local Similarity 81.7%; Pred. No. 0;  
Matches 2512; Conservative 0; Mismatches 526; Indels 37; Gaps 6;  
QY 201 GTTCTCTCAGGATGAGCACAGCTGGGAGATGGCAGTGTGTAATGGCTGCGCCAGG 260  
DB 54 GCGCTCCGCGGTGAGCAGCAGCTGGGAGATGGCCGTGTGTAATGGCTGCGCCAGG 113  
QY 261 TGCCCTTGGCTCAGAGTGGGCGACAGAGAACTCCCCAGCCAAAGGAGAGAGCGCC 320  
DB 114 CACTCTCGGTCAGCAGTGGGCGACAGAGAGGCTTGCACCAAGCAAGGAGAGACACA 173  
QY 321 GCCACTGGGGAAGAACAGAGCTCCGTCTACAGCTTGGAGCGCTGGAGAGAGAGCCCTGT 380  
DB 174 GTCACTGGGGAAGAGAGAGCTGATCTTCAAGCTTGGAGCGCTGGAGAGAGAGCCCTGT 233



1314 AGGCTCTTTCGGCGAGGTCACAGAAATGAAGACAAGCAGACAGGCTTCCAGTGTGCTGT 1373  
1515 CAAAAGGTTCGGCTTGGGAAGTATTTTCGGGCGAGAGGAGCTGATGGCATGTGCAGGATTGAC 1574  
1374 CAAAAGGTTCGAGCTCGAGGTTTTCGGGTAGAGGAACCTAGTGGCTGTGCTGCTGAG 1433  
1575 CTCACCCAGAAATTTGCCCTTTGTATGGAGCTGTGAGAGAAGGCTTGGGTCAACATCTT 1634  
1434 CTCGCCAGAAATTCGTTCCCTCTCTATGGAGCTGTGAGAGAAGGCCCGTGGGTGAACATCTT 1493  
1635 CATGGAGCTGTGGAAGGTGCTCCCTGGGCCAGCTGCTCAAGGAGCAGGCTGTCTCCC 1694  
1494 CATGNACTGTAGAAAGTGGCTCCTGGGTGAGCTCATAAAGCAAAATGGGCTGTCTGCC 1553  
1695 AGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTGGAATACCTCCACTC 1754  
1554 AGAAGCCGAGCCCTTTACTACTCTGGGCCAGGCCCTGGAGGGCTGGAGTACTCCACAC 1613  
1755 ACGAAGGATTCGATGGGGAGCTCAAAAGCTGACAACTGCTCTCTGTCACGCGATGGGAG 1814  
1614 ACGCAGGATTCGATGGCGATGTCAAAAGCTGACAACTGCTCTCTGTCACGATGGAAG 1673  
1815 CCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACTGATGCTGGGAA 1874  
1674 CCGAGCGCCCTCTGCGACTTTGGCCAGCCCTTGTGCTGCAACTGACGCCCTTAGGGAA 1733  
1875 GTCTTGTCTACAGGGGACTACATCCCTGGCACAGAGACCCACATGGCTCCCGAGGTGGT 1934  
1734 ATCTTGTCTACAGGGGACTACATCTCTGGCACGGAGACCCACATGGCACCAAGTGGT 1793  
1935 GCTGGCAGGAGCTCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGCTGCA 1994  
1794 GATGGAAAGCCCTCGCATGCGAAGTGGACATCTGGAGCAGCTGCTGATGATGCTCCA 1853  
1995 CATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCTGCTCAAGAT 2054  
1854 CATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGCTCTGCTCAAGAT 1913  
2055 TGCCAGCAGGCTTCGGCTGTGAGGGAGATCCACCCCTCTCGCGCCCTCTCAAGGCCA 2114  
1914 TGCCAGCAGGCTTCACCGATCAGGGAGATCCACCTCTCTCGGCAACCCCTCAAGGCCA 1973  
2115 GGCCATCAAGAGGGGCTGAGGAAGAGCCCATCCACGCGTGTCTGAGCGGAGCTGG 2174  
1974 GGCCATCAAGAGGGGCTGAGGAAGAGCCCGTCCACCGAGCATCTGCAATGGAGCTTCG 2033  
2175 AGGNAAGTGAACCGGGCACTACAGCAAGTGGAGGCTTGAAGAGCCCTTGGAGGGAGA 2234  
2034 GAGGAAGTGGGCAAGGCATACAGGAAGTGGAGGCTTGAAGAGCCCTTGGAAAGGAGA 2093  
2235 ATATAAGAACCAAGACATCCACCGCCAAATCAAGGCCAATTACCACAGAGCCCTCCATGC 2294  
2094 ATATAAGAACCAAGACCTCCACCAAGACCAAGCCACCTGCCACAGACCTTACCTAC 2153  
2295 CCAGCCGAGAGGCTTTTCGCCAAGGGCCCGAGGCGCCGAGCTGAGGAGACCAAGG 2354  
2154 TCCGCCGAGAGA-----GAAACCCACAGCCCAAGGCCAACAACAGA 2192  
2355 CAGAGCCCTTAAGCTCAGGCTCTCTCCCAACAGAGCCCGCCAGAGCCCAACAAGTCTCC 2414  
2193 CGGGGCTCTGAGGCTCAGGCTCTCTACCGCCAGAACCCACAGAACCGAGCAAGGCC 2252  
2415 TCCCTTGACTTTGAGCAGGAGGAGTCTGGGATGCTGGGAACCCCTTACTCTGCTCCCT 2474  
2253 AGCCCTGAACCTGAGCAAGGAGGAGTCTGGCACAATGGGAACCCCTGCTCTGCTCCCT 2312  
2475 GGAGCCAGCCCTTCAGAAAACCCAGCTCACAGAGCGGGAAGCAACCGTCCCGGAGCA 2534  
2313 GGACCCAGCACTGCCAAGGCCCGCAGCTTCCAGACCGGAGGGCAACCTTGGCAGAGCT 2372  
2535 GGAAGTCTGAGGCTGGAAATAGAAATATTCCTCAACAGCCTGTCCAGGCAATTTCTCT 2594

Db 2373 GGAGCTACAGCAACTGGAGATAGAACTGTTTCTCAACAGCCTGTCCAGCGGTTCTCTCT 2432  
Qy 2595 GGAGGAGCAGAGCAAAATTTCTCGTGCCTCAGCATCGACAGCCTCTCTCTGTGCGATGA 2654  
Db 2433 GGAGGAACAGGAACAAATCTCTCTCTCTCTCAGCATCGACAGCCTCTCGTGTGATGA 2492  
Qy 2655 CAGTGAGAAGAACCCATCAAAAGGCTCTCAAGGCTCGGGGACACCCCTGAGCTCAGGGCT 2714  
Db 2493 CAGTGAAGAAATCATCGAAGGCTCTCAGAGCTCAGGGACACCCCTGAGTTCTGGGCT 2552  
Qy 2715 ACATCTCTGGAGCAGCAGGCCGAGGCTCGAAGCTCAGCTTGGAAACATGGTGTGGGCCG 2774  
Db 2553 GCATCTTGGAAACAGCCAGCTGAGGCAAGAACTGAGCTGACAGCGGCGCTGGGCCG 2612  
Qy 2775 GGSGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAAATACAGTCTCT 2834  
Db 2613 GGSGCGGCTTACTGACATCCGAGCTTCAACGGGTCAGGTTCCAGATCCAGTCTCT 2672  
Qy 2835 TAATGGTGAACACTCTGACATCCGGGAGTTCCACCGGTCAAAGTGGGAGACATCGCCAC 2894  
Db 2673 CAATGGCGAAACACTGTCATATCCGGGAATTCACCGGTCAGGTTGGGAGACATTGCCAC 2732  
Qy 2895 TGGCATCAGCAGCAGATCCAGCTGCAGCTTCAGCTTGGTCAACAAAGACGGGCGCC 2954  
Db 2733 CGGCATCAGCAGCAGATCCAGCCACAGCTTTTTCAGCTGTGTGACAAAGATGACAGCC 2792  
Qy 2955 TGTTCGTACGACATGAGAGTGGCAGACTCGGCGATCGACTGCAGTGCACACTGGGCC 3014  
Db 2793 TGTTCGTATGACATGAGGTCGCGAGCTCGGCGATCGACCTGCAGTGCACCTGGGCC 2852  
Qy 3015 TGATGGCAGCTTCGCTGGAGCTGGAGGTCGAGCTCAAGCTGGCCAGCTGGAGAAACAGGCCCTA 3074  
Db 2853 TGATGGCAGCTTTGCTTGGACCTGGAGGTCGAGCTCAAGCTGGTCAAGTGGAGAACCGACCTA 2912  
Qy 3075 ACCCTGCCCTCCACCGCGGCTCCACACTG---CCGGAAGCAGCCTTCTCTGCTCGGTGC 3131  
Db 2913 GTCTGTCTTCCACACAGCTCCACTCTGCGCACGGGAGCTTCTTTGATCTCAGTGC 2972  
Qy 3132 ACATGTCTGCCCTGAAACACAGGCTCAGCGCTTCCAGGGGATTCGACGCCCCCGGCT 3191  
Db 2973 TCCTGTGTCTGAGATGCTGG---ATGAGTCTGCACAGGACCTGCCAGTCCCTGGTT 3029  
Qy 3192 CACAGTGGGAACCAAGGCTTCGACAGCAAGTGGGGGCA----AGCAGAAATGCTTCCC 3247  
Db 3030 CAGCATGTAGACAGGCTTGTGAGCAGCAGCAAGGGAACAGAGTAAAGCACCTCCC 3089  
Qy 3248 AGGATTCACACCTG 3262  
Db 3090 AGGACTTGCCTCTG 3104

RESULT 4  
AY413578  
LOCUS  
DEFINITION Pan troglodytes MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY413578  
VERSION AY413578.1 GI:39769540  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
REFERENCE 1 (bases 1 to 2765)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarial,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 2765)



```
Qy 2171 TGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCTTGGAGGG 2230
Db 1861 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 2231 GAGATATAAGAACCAAGACATCCACCGCCAATCAAGCCAAATTACCAACAGACCCCTCC 2290
Db 1921 GAGATATAAGAACCAAGACATCCACCGCCAATCAAGCCAAATTATACCAAGACCCCTCC 1980
Qy 2291 ATGCCACCGCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCCGGCCAGCTGAGGAGACAA 2350
Db 1981 ATGCCACCGCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCCGGCCAGCTGAGGAGACAA 2040
Qy 2351 CAGCAGAGCCCTTAAGCTCAGCCTCTCTCCACCCAGAGCCCCCAGAGCCCAACCAAGT 2410
Db 2041 CAGCAGAGCCCTTAAGCTCAGCCTCTCTCCACCCAGAGCCCCCAGAGCCCAACCAAGT 2100
Qy 2411 CTCTCTCTCTGACTTTCAGCAAGAGGAGCTGGGATGTGGGACCCCTTACCTCTGTCTCT 2470
Db 2101 NTCTCTCTCTGACTTTCAGCAAGAGGAGCTGGGATGTGGGACCCCTTACCTCTGTCTCT 2160
Qy 2471 CCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAAGCAACGTCCTCCGG 2530
Db 2161 CCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACAGAGCGGAAAGCAACCTTCCCGG 2220
Qy 2531 AGCAGGAAGTGCAGAGCTGGAATAGAAATTATCTCTCAACAGCCTGTCCAGCAATTTT 2590
Db 2221 AGCAGGAAGTGCAGAGCTGGAATAGAAATTATCTCTCAACAGCCTGTCCAGCAATTTT 2280
Qy 2591 CTCTGGAGGAGCAGGACCAATTTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTCTGG 2650
Db 2281 CTCTGGAGGAGCAGGACCAATTTCTCTGTCCTCAGCATCGACAGCTCTCCCTGTCTGG 2340
Qy 2651 ATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACACCTGAGCTCAG 2710
Db 2341 ATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACACCTGAGCTCAG 2400
Qy 2711 GCCTACACTCTCTGGAGCAGCCAGCCGAGGCTCGAAGCTCCAGCTGGAAACATGCTGTGG 2770
Db 2401 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 2771 CCGGGGGGGCCACCGACACCCCAAGCTATTTCATGTGTGAAGTCAAAATACAGT 2830
Db 2461 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 2831 CTCTTAATGTGAACCTGCACATCCGGAGTTCCACCGGTCCAAAGTGGGAGACATCG 2890
Db 2521 CTCTTAATGTGAACCTGCACATCCGGAGTTCCACCGGTCCAAAGTGGGAGACATCG 2580
Qy 2891 CCACTGGCATCAGCAGCAGATCCAGCTGAGCCTTTCAGCTTGGTCAACCAAGACCGGC 2950
Db 2581 CCACTGGCATCAGCAGCAGATCCAGCTGAGCCTTTCAGCTTGGTCAACCAAGACCGGC 2640
Qy 2951 AGCTGTTCGTACGACATGAGAGTGCAGACTCGGGATCGACTGAGTGACACTGG 3010
Db 2641 AGCTGTTCGTACGACATGAGAGTGCAGACTCGGGATCGACTGAGTGACACTGG 2700
Qy 3011 CCCCTGATGCGAGCTTCGCTGGAGTGGAGGTCAGAGTGGCAGCTGAGGACAGGC 3070
Db 2701 CCCCTGATGCGAGCTTCGCTGGAGTGGAGGTCAGAGTGGCAGCTGAGGATAGGC 2760
Qy 3071 CCTAA 3075
Db 2761 CCTAA 2765
```

```
RESULT 5
AV413579 2829 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION Mus musculus MAP3K14 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AV413579
VERSION AV413579.1 GI:39769541
```

```
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2829)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2829)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..2829
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2829
/genes="MAP3K14"
/locus_tag="HCM4937"
gene
ORIGIN
Query Match 42.1%; Score 1942; DB 9; Length 2829;
Best Local Similarity 79.0%; Pred. No. 0;
Matches 2251; Conservative 0; Mismatches 571; Indels 27; Gaps 3;
```

```
Qy 232 ATGGCAGTGTGTAATGGCTTCCAGAGTGGCTTGGCTCAGCAGTGGGCGAGCAGAAG 291
Db 1 ATGGCGTGTGTAAGTGGCTTCCCGGCACTCTCTGGTTCAGCAGTGGGCGAGCAGAAG 60
Qy 292 GAATCTCCCAAGCAAGGAGAGAGCCGCCACTGGGGAGAGAAACAGAGTCCGTCTAC 351
Db 61 GAGCTTCCCAAGCAAGGAGAGAGACAGTCACTGGGGAAGAGCAGAGTGCATCTTC 120
Qy 352 AAGCTTGAAGCCCTGGAGAGAGCCCTGTCTTCTCGGGAAGTGGGAGATCCTGAATGAC 411
Db 121 AAGCTTGAAGCCCTGGAGAGAGCCCGTGTCTTGTGGGAAGTGGGAGATCCTGAACGAC 180
Qy 412 GTGATTACCAAGGCACAGCAAGGAGGCTCCGAGGAGGCGGCGAGCTGCTATC 471
Db 181 GTGATCAACCAAGGCACAGCAAGGAGGCTCTGAGGAGGAGCACCGGCCATCTCCATC 240
Qy 472 ATCGCCAGAGCTGAGTGTGAGAAATAGCAAGAGTTACGCCCCACCTTTTCAGAACGCAAT 531
Db 241 ATCGCCAGAGCTGAAATGTGAGAAATAGCAAGAGTTACGCCCCACCTTTCTCAGAGCGCAAT 300
Qy 532 TTCTATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 591
Db 301 TTCTATCGCGGGCTCACAGCAGTACAGCCAGTCTCTGATCAAAATCCCCCAACAAT 360
Qy 592 GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGGTGGAGGAGAAACGCTCGCAGC 651
Db 361 GTGGCCCATCTCAACTGAAGCAAAATGGCCCGTGTGTGGGAGGAGAAACGTCACGGC 420
Qy 652 AAAGCCCGGAAGAAACGGAAGAGAGTCAAAAGTCCCTGGTCTATGACGAGTGGCC 711
Db 421 AAAGCCCGGAAGAAACGTAAGAGAGAGTCAAAAGTCACTGGGCCAGGAGGAGTGGCC 480
Qy 712 TTGGCCAAACCCCTCCCGCAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTCCAGGAG 771
Db 481 TTAGCCAAAGCCCTCGCCAGAACCCCTGAGCAAGAGAGTGTATCCATCCCAAGTACAGAA 540
```

QY	772	GATGAGTCTCCACTCGGCGCCCATATGTTAGAAAACACCCCGAGTTCCACAGCCTCTG	831
Db	541	GATGAGTCTCCACTAGGCAACCTCTATGCCAGAAATGTTCTCCAGTTTCCACCAAGCCTCTG	600
QY	832	AAGGAACCAAGGCTTGGGCAACTCTGTTTTAAGCAGCTTGGGAGGCGCTAAGGCGGCT	891
Db	601	GGGGACCAAGGCTTGGCCACCTGTCTTTAAGAACAAAGATGAAGGCTTGGACCGGTA	660
QY	892	CTGCCCTCGATCAGAACTCCACAAACTGATCAGCCCTTGGAAATGTTGAAACCAAGCTGTGG	951
Db	661	CTGCCCTCGACCAAACTCCACAAACTGATCAGCCCTTGGAAATGTTGAAACCAAGCTGTGG	720
QY	952	AAACTCAGACACCCAGAGAGGAGGCGCCCTGCGCCCTGCCAGCGACCCCTTCCCTTAT	1011
Db	721	AAACTCAGACACCCAGAGGAGGCGCCCTGCGCCCTGCCAGCGACCCCTTCCCTTAT	780
QY	1012	AGCAGACTCCCTCATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAAACCTCACCCCTCTG	1071
Db	781	AGCGGAATGCCCATCTCTTCCCATTTCTACCCCTTGAGGCCCTTGAAACCTTATGCTG	840
QY	1072	GAGT---CCTTCTGGGAAACTGGCCTGTGTAGACAGCAGAAACCTTGCCTGACCCCA	1128
Db	841	GACTCTGCGCTCTGGACAACTAGCCGTGTGAGCGGCCAGCGCCTCTGCCTGGCCCA	900
QY	1129	---CACCTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGGCCACACCTG	1185
Db	901	CGGCATCTAAGCCNACTGGCCCATGGAGACAGTCAAGAGCCCTGCTGGGCCACACCTG	960
QY	1186	GAGCCAGCTGCTGTCTGCTGGTCCCATGAGAAATTTTGTGGAGAAATACCTAGTG	1245
Db	961	GAGTCCAGCTGCGCGTCTCGGGTGCCTTAGAAAAGTTTCCCGTGGAGGAATACCTGGTG	1020
QY	1246	CATGCTCTCAGGAGAGCTGAGCTCAAGCCAGGCGCCACAGCCTGACAGCCTTGGCCAG	1305
Db	1021	CATGGCTCTAAGGAAGTGTAGCTCAGCCAGGCGCCACAGCCTGCGCAGCCTTGGCTAAG	1080
QY	1306	ACCTGGGAGCAGCGGGCTTCAGATCCCGGGAGCGCCAGCCCAAACTGAGGACAAACGAG	1365
Db	1081	ACATGGTCTCTGGGAAGCCCAAGCTGCAGAGGCTCGGCCCGAACTGAGGACAAACGAG	1140
QY	1366	GGTGTCTGCTCACTGAGAACTCAAGCAGTGGATTATGATGACCGGAGAAAGTCCAC	1425
Db	1141	GGGGTCTCTTACTGAGAACTCAAGCCAGTGGATTATGATGATCGAGAGAGGTCCAC	1200
QY	1426	TGGGCCACGACACCTCGCTCGGCTGGGAGAGGCTCCTTCGGAGAGGTGCACAGATGGAG	1485
Db	1201	TGATGACACACACCTCGGCTGGGCTGAGGCTCCTTCGGGCGGTTCCACAGAAATGAAG	1260
QY	1486	GACAGCAGACTGGCTTCCAGTGCCTGTCAAAAAGGTGGCGCTGGAAAGTATTTTCGGCA	1545
Db	1261	GACAGCAGACAGGCTTCCAGTGTCTGTCAAAAAGGTACGACTCGAGGTGTTTCGGTA	1320
QY	1546	GAGGAGCTGATGGCATGTGCAGATTGACTCACCCAGAAATGTCCTTTGTATGAGCT	1605
Db	1321	GAGGAACTAGTGGCTGTGCTGGTCTGAGCTCGCCAGAAATCGTCCCTCTCTATGAGCT	1380
QY	1606	GTGAGAGAGGCTTGGGTCAACATCTTCATGAGCTGCTGGAAGGTGGCTCCCTGGGC	1665
Db	1381	GTGAGAGAGGCTTGGGTGAACATCTTCATGAACTGCTAGAAGGTGGCTCGCTGGGT	1440
QY	1666	CAGCTGGTCAAGGACAGGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCGGGCCAG	1725
Db	1441	CAGCTCATAAAGCAAAATGGGCTGTCTGCCAGAAAGCCGAGCCCTTTACTACTCGGGCCAG	1500
QY	1726	GCCCTGGAGGCTCTGGAATACCTCACTCAGAAAGATTCTGCATGGGGAAGTCAAAGCT	1785
Db	1501	GCCCTGGAGGCTGGAAGTACCTCCACACACGAGGATTCTGCATGGCGATGTCAAAGCT	1560
QY	1786	GACAACTGCTCTCTCAGCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCT	1845
Db	1561	GACAACTGCTCTCTCAGTGTGAAGCCGAGCGGCGCTCTGCGACTTGGCCACGCCC	1620

QY	1846	GTGCTCTTCAACCTGATGGGCTGGAAAGTCTTGTCTCACAGGAGCTACATCTCCTGGC	1905
Db	1621	TTGTGCTTGCACCTGACGBCCTTAGGGAAATCTTGTCTCACAGGAGCTACATCTCCTGGC	1680
QY	1906	ACAGAGACCCACATCGCTCCGAGGTGTGTGGCAGAGCTGCGAAGCTGGAATGAT	1965
Db	1681	ACGAGAGACCCACATGGCACCAAGAGTGTGTGGAAAGCCCTGCGATGCGAAGGTGAC	1740
QY	1966	GTCTGGAGAGCTGTGTATGATGTGCATGTCAAAGCTGCGAAGCTGCGCTGTGAGGAGATC	2025
Db	1741	ATCTGGAGCAGCTGTGATGATGTCCACATGTCTCAACGCTGCGAAGCTGCGCTGTGAG	1800
QY	2026	TTCTTCCAGAGGCGCTCTGCTCAAGATGCGCAGGAGCTTCCGCTGTGAGGAGATC	2085
Db	1801	TACTTCCAGAGGCGCTTGTCTCAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1860
QY	2086	CCACCTCTCTGCGCCCTCTCACAGCCAGGCCATCAAGAGGGCTGAGGAAAGAGGCC	2145
Db	1861	NN	1920
QY	2146	ATCCACCCGCTGTCTGACGCGAGCTGGAGGGAGAGTGAAACCGGGCACTACAGCAAGTG	2205
Db	1921	NNNTG	1980
QY	2206	GGAGTCTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAT	2265
Db	1981	GGAGTCTGAAGAGCCCTTGGAAAGGAGAAATATAAGAAACCAAGACCTCCACCCAGAC	2040
QY	2266	CAAGCAATTAACCAAGACCCCTCATGCCAGCGCAGAGAGCTTTGCCAAGGGCCCCCA	2325
Db	2041	CAAGCACTTCCACAGACCCCTACTACTCGCGCAGAGA-----	2081
QY	2326	GGGCCCCGGCAGCTGAGGAGACAAACAGGAGAGGCCCTTAAGCTCCAGCTCTCTCCCA	2385
Db	2082	--GAACCCACAGCCCAAGGCCAACACAGAGGGGCTCTGAGCCTCAGCCTCTCTACCG	2139
QY	2386	CCAGAGCCCCCAGAGCCAAACAAAGTCTCCTCCCTTGTACTTTGAGCAAGAGGAGTCTGG	2445
Db	2140	CCAGAACCAACAGAACCGAGCAAGAGGCCAGCCCTGAACTGAGCAAGAGAGGATCTGGC	2199
QY	2446	ATGTGGGAACCCCTTACCTCTCTCTCTCTGAGCAGCCCTCTGCGAGAAACCCAGCTCA	2505
Db	2200	ACATGGGAACCCCTGCTCTCTCTCTCTGAGCCAGCCACTGCGCAAGGGCCCCAGCTTC	2259
QY	2506	CCAGGCGGAAGCAACCGTCCCGGAGCAGAACTGCGAGCAGCTGGAATAGAAATATTC	2565
Db	2260	CCAGACCGGGCAACCTTGCCAGAGCTGAGCTACAGCAACTGGAGATAGAACTGTGT	2319
QY	2566	CTCAACAGCTGTCTCCAGCCATTTTCTGAGGAGCAGGAGCAAAATCTCTCGTGCCTC	2625
Db	2320	CTCAACAGCTGTCTCCAGCCGTTCTCTGAGGAAACAGGAAACAAATCTCTCTGCTC	2379
QY	2626	AGCATCGACAGCT	2685
Db	2380	AGCATCGACAGCT	2439
QY	2686	AGCTCGGGGACACCTGAGCTCAGGCTACCTCTCTGAGCAGCCAGCGGCGAGCTCGA	2745
Db	2440	AGCTCAAGGAGACCTGAGTCTGGGCTGACCTCTTGGAAACAGCCAGCTGAGGCAAGA	2499
QY	2746	AGCTCCAGCTGGAAATCTGCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTTC	2805
Db	2500	ACCTGAGCTGACAGCGGCTGCGCCCGGGGGCGGCTACTGTACATCCCGAGCTACTTC	2559
QY	2806	AATGCTGAAGTCCAAATACAGTCTCTTAACTGTTGAAACCTGTGCACTCCGGGAGTTC	2865
Db	2560	AACGGGTCAAGTCCAGATCCAGTCTCTAAATGGCGAAACACCTGCAATATCCGGGAATTC	2619
QY	2866	CACCGGTCAAAGTGGGAGACATCGCACCTGGCATCAGCAGCCAGATCCCAAGCTGACGCC	2925
Db	2620	CACCGGTCAAAGTGGGAGACATTTGCCACCGGCATCAGCAGCCAGATCCCAAGCAGCT	2679
QY	2926	TTACGCTTGTTCACCAAGAGCGGCGAGCCTGTTCTGCTACGACATGGAGGTGCCAGACTCG	2985

```
Db      2680 TTACGCTGGTGACCAAGATGGACAGCCTGTTGCTATGACATGGAGGTGCCAGACTCG 2739
Qy      2986 GGCATGACCTGACGTGCACACTGGCCCTGATGCGACGCTTTCGCTGAGCTGGAGGTC 3045
Db      2740 GGCATGACCTGACGTGCACCTGGCCCTGATGGCAGCTTGTGTTGACCTGGAGGTC 2799
Qy      3046 AAGCATGCCAGCTGGAGAACAGGCCCTA 3074
Db      2800 AAGCATGCTCAGCTGGAGAACGACCTTA 2828

RESULT 6
LOCUS   CR612342                1640 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODK011Y14 of HeLa cells Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR612342
VERSION   CR612342.1 GI:50493149
KEYWORDS HTC; CNSLT cDNA.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1640)
AUTHORS   Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL   Full-length cDNA libraries and normalization
          Unpublished
REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue
          Gaithersburg, MD 20878
          (bases 1 to 1640)
REFERENCE 2. (bases 1 to 1640)
AUTHORS   Genoscope.
JOURNAL   Direct Submission
          Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
          - Web : www.genoscope.cns.fr)
COMMENT   1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of invitrogen.
FEATURES             Location/Qualifiers
     source           1..1640
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CSODK011Y14"
                     /tissue type="HeLa cells Cot 25-normalized"
                     /plasmid="pCMVSPORT_6"

ORIGIN
Query Match      33.1%; Score 1521.2; DB 3; Length 1640;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 1615; Conservative 0; Mismatches 23; Indels 10; Gaps 7;

Qy      2922 AGCTTCAGCTTGGTACCAAGAGCGGACGCTGTTGCTGACATGGAGGTGCCAGA 2981
Db      1 AGCTTCAGCTTGGTACCAAGAGCGGACGCTGTTGCTGACATGGAGGTGCCAGA 60

Qy      2982 CTCGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTTCGCTGGAGCTGGAG 3041
Db      61 CTCGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGCAGCTTTCGCTGGAGCTGGAG 120

Qy      3042 GGTCAAGCATGGCAGCTGGAGAACAGGCCCTTAACCTTGCCCTCACGCCGCTCCACA 3101
Db      121 GGTCAAGCATGGCAGCTGGAGAACAGGCCCTTAACCTTGCCCTCACGCCGCTCCACA 180

Qy      3102 CTCGGGAAGACGCTTCTGCTGGTGACGATGCTGCCCTGAAACACAGGCTCAGC 3161
Db      181 CTCGGG-AAGACGCTTCTGCTGGTGACGATGCTGCCCTG-AAACACAGGCTCAGC 238

Qy      3162 CGTTCCAGGGGATTGCCAGCCCCCGGCTCACAGTGGGAACGAGGGCTCGCAGCAGCA 3221
```

```
Db      239 CGTTCCAGGGGATCTGCCAG-CCCCCGGCTCAGCAGTGGGACCGAGGGCTTCGACAGCA 297
Qy      3222 AGGTGGGGGAAGCAGAAATGCTCCCAAGATTTCACACCTGAGCCCTGCCCCACCCCTGCT 3281
Db      298 AGGTGGGGGAGCAGAAATGCTCCCAAGATTTCACACCTGAGCCCTGCCCCACCCCTGCT 357
Qy      3282 GAAAAAACA-TCCGCCACAGTGAAGAGACAGAAAGAGGATGGCAGGAGTTACTCTGGGAAA 3340
Db      358 GAGAAAAACACTCCGCCACGTGAAGAGACAG-AGGAGGATGGCAGGAGTTACTCTGGGAAA 416
Qy      3341 CAAAAACAGGATCTTTTCTGCCCCCTGCTCAGTCGAGTTGGCCCTGACCCCTTGGATCA 3400
Db      417 CAAAACAGG----ATCTTCTCTGCCCTGCTCAGTCGAGTTGGCCCTGACCCCTTGGATCA 472
Qy      3401 GTGACCAATTTGTCGACAGACAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGGGGTGGGC 3460
Db      473 GTGACCAATTTGTCGACAGACAGGGGAGAGCAGCTTCCAGCCTGGGTCAAGAGGGGTGGGC 532
Qy      3461 GAGCCCTTCGGCCCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCA 3520
Db      533 GAGCCCTTCGGCCCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCA 592
Qy      3521 AACTCAGGTTTCAGTGCAGAACACAGGTCAGCAGGTATGCCCCCGTAGGTAAAGGGGCC 3580
Db      593 AACTCAGGTTTCAGTGCAGAACACAGGTCAGCAGGTATGCCCCCGTAGGTAAAGGGGCC 652
Qy      3581 CTCTAAACCCCTTGGCTGGGCTCACCTGGCCAGCTCACCCCTTTTGGGTAGGGGAAA 3640
Db      653 CTCTAAACCCCTTGGCTGGGCTCACCTGGCCAGCTCACCCCTTTTGGGTAGGGGAAA 712
Qy      3641 GAATGCTGACCCCTGGGAAGGCTCCCTGGTAGAATAACACACACTTTTCAGGTGTGTTGCA 3700
Db      713 GAATGCTGACCCCTGGGAAGGCTCCCTGGTAGAATAACACACACTTTTCAGGTGTGTTGCA 772
Qy      3701 ACACAGTCTCTGAGTTGACCTCTGGTTACGCAAGGACCAAGAGGTGTGTAAGTGAAG 3760
Db      773 ACACAGTCTCTGAGTTGACCTCTGGTTACGCAAGGACCAAGAGGTGTGTAAGTGAAG 832
Qy      3761 TGGTTCAGTCCCGACACATGSCCCTTTGCTGCTGGCTACCACTCTTCCCCAGAGCA 3820
Db      833 TGGTTCAGTCCCGACACATGSCCCTTTGCTGCTGGCTACCACTCTTCCCCAGAGCA 892
Qy      3821 GCAGGCCCGAGCCCTTCAGGGCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCT 3880
Db      893 GCAGGCCCGAGCCCTTCAGGGCCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCT 952
Qy      3881 CATCTGTAAAGGTGAAGGTTGATGACGATATGCTGACAGGAACAGTCTGTGATGAGAC 3940
Db      953 CATCTGTAAAGGTGAAGGTTGATGACGATATGCTGACAGGAACAGTCTGTGATGAGAC 1012
Qy      3941 ATGATCAGTCTAAGGAAGCAGCAGAGAGAGAGCG-TCCGGCGCCCGAGCCCACTATCA 3999
Db      1013 ATGATCAGTCTAAGGAAGCAGCAGAGAGAGAGCGTCCGGCGCCCGAGCCCACTATCA 1072
Qy      4000 GTGTCCAGCGTCTGGTTTCCCCAGAGCAGCTCAGCATACACTGACACTCACCTGCC 4059
Db      1073 GTGTCCAGCGTCTGGTTTCCCCAGAGCAGCTCAGCATACACTGACACTCACCTGCC 1132
Qy      4060 CTGCCCCCTGGCCAGAGGTTACTGCCGACGGCAGCTTTGCACTCTGATGACCTCAAGCACT 4119
Db      1133 CTGCCCCCTGGCCAGAGGTTACTGCCGACGGCAGCTTTGCACTCTGATGACCTCAAGCACT 1192
Qy      4120 TTCAATGCTGCCCTCTGCGAGGGCAGGGCAGGTGACACTGTAGGAGCATAGCAGC 4179
Db      1193 TTCAATGCTGCCCTCTGCGAGGGCAGGGCAGGTGACACTGTAGGAGCATAGCAGC 1252
Qy      4180 CAGGAGATGGGTGAAGGGACACAGTCTTGAGCTGTCCACATGATGTGACTCTCTCAAC 4239
Db      1253 CAGGAGATGGGTGAAGGGACACAGTCTTGAGCTGTCCACATGATGTGACTCTCTCAAC 1312
Qy      4240 CTCTTCCAGATTTCTCTAAGAATAAGCACCCCTTTCCCAATTCGCCAGCTTAGCCTCTTC 4299
```

```
Db 1313 CTCTTCAGAGTTCTTCTAAGAAATAGCACACCCCTTCCCCATTGCCCCAGCTTACGCTCTTC 1372
QY TCCACAGGGAGTACTCAGACTCAGTACATTAATCAGCTGTGAATCGTCAGGGGT 4359
Db 1373 TCCACAGGGAGTACTCAGACTCAGTACATTAATCAGCTGTGAATCGTCAGGGGT 1432
QY GTCTGTAGCTCAACCTCTCTGGGGCAGGGAGCGCCGAGACTCCGTGGGAGAAGCTCAT 4419
Db 1433 GTCTGTAGCTCAACCTCTCTGGGGCAGGGAGCGCCGAGACTCCGTGGGAGAAGCTCAT 1492
QY CCCACATCTTGGCAAGACAGCCCTTGTCCAGCTGTCCATTTAGTCAGACTGCTCCCG 4479
Db 1493 CCCACATCTTGGCAAGACAGCCCTTGTCCAGCTGTCCATTTAGTCAGACTGCTCCCG 1552
QY GGAGAGAGCCCCGGCCCCCAGACACATAAAGAACTGTCAGCTTGGTACTCGAGAGTCTGG 4539
Db 1553 GGAGAGAGCCCCGGCCCCCAGACACATAAAGAACTGTCAGCTTGGTACTCGAGAGTCTGG 1612
QY TTGTAGAGAAGCTCTTTTGAAGCAATAA 4567
Db 1613 TTGTAGAGAAGCTCTTTTGAAGCAATAA 1640
```

## RESULT 7

```
AL578455/c
LOCUS AL578455 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens 1095 bp mRNA linear EST 07-APR-2004
DEFINITION CDNA clone CS0DK011YML4 3-PRIME, mRNA sequence.
ACCESSION AL578455
VERSION AL578455.3 GI:46257344
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:31316669.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
8426.r
```

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DK011BG07NP1&c=8426.r.

## FEATURES

```
source
1..1095
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK011YML4"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

## ORIGIN

```
Query Match 20.5%; Score 943; DB 1; Length 1095;
Best Local Similarity 94.1%; Pred. No. 2.9e-226;
Matches 1016; Conservative 13; Mismatches 44; Indels 7; Gaps 6;
QY 3490 GCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAACTCAGTTTCAGTGCAGAACCAAGTCTAG 3549
```

```
Db 1075 GGTKVTGTARATKKCAAGTGTAAAGCCCAAACTCAGTTTCAGTGCAGAACCAAGTCTAG 1016
QY 3550 CAGGTATGCCGCCCGTGTAGTTTAAAGGGGCCCTCTTAAACCCCTTGGCTGGCCCTCAGCTGG 3609
Db 1015 MA-GKATGCCGCCCGTGTAGTTTAAAGGGG--CCYTTTAAACCCCTTGGCTGGCCCTCAGCTGG 959
QY 3610 CCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCCTTGCACCCCTGGGAAGGCTCCCTGG 3669
Db 958 CCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATG-CTGACCCCTGGGAAGGCTCCCTGG 900
QY 3670 TAGAATAACACCACTTTTTCAGGTTTGTGCAACACACAGTCTCTGAGTTGACCTCTGTTTCA 3729
Db 899 TAGAATAACACCACTTTTTCAGGTTTGTGCAACACACAGTCTCTGAGTTGACCTCTGTTTCA 840
QY 3730 GCCAAGGACCAAGAGGCTGTGTAAGTGAAGTGTCTTCACTCCCAAGACATGTGCCCT 3789
Db 839 GCCAAGGACCAAGAGGCTGTGTAAGTGAAGTGTCTTCACTCCCAAGACATGTGCCCT 780
QY 3790 TTGCTGTCTGGCTTACCACTCTTCCCAAGACAGCAGGCCCCGAGCCCTTTCAGGCCCCAGCA 3849
Db 779 TTGCTGTCTGGCTTACCACTCTTCCCAAGACAGCAGGCCCCGAGCCCTTTCAGGCCCCAGCA 720
QY 3850 CTGCCCCAGA-CTGCTGGCACTCAGTTCCCTCATCTGTATAAGGTGAAGGGTGTGCAAG 3908
Db 719 CTGCCCCAGAGCTGAGTGTGTAAGTGAAGTGTGTAAGGTGAAGGGTGTGCAAG 660
QY 3909 ATATGCTGTGACAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTTAAGAAAGCAGCAG 3968
Db 659 ATATGCTGTGACAGGAAACAGTCTGTGGATGGACATGATCAGTGTCTTAAGAAAGCAGCAG 600
QY 3969 AGAGACG--TCCGGGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTCCCAAGACGA 4027
Db 599 AGAGACGCTCGCGGCCCCAGCCCCCACTATCAGTGTCCAGCGTGTGTTCCCAAGACGA 540
QY 4028 CAGTCTCAGCATCAGACTCAGACTACCTGCTGCCCTGCTGCCAGAGGGTACTGCCGAC 4087
Db 539 CAGTCTCAGCATNANACTGANACTCACCTGCTGCCCTGCTGCCAGAGGGNACTGCCGAC 480
QY 4088 GGCACCTTGGCACTCTGATGACCTCAAGACACTTTTCATGGCTGCCCTCTGGCAGGGCAGGG 4147
Db 479 GGCACCTTGGCACTCTGATGACCTCAAGMACCTTTTCATGGCTGCCCTCTGGCAGGGCAGGG 420
QY 4148 CAGGCACTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACAGTCT 4207
Db 419 TAGGSCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACAGTCT 360
QY 4208 TGAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4267
Db 359 TGAGCTGTCCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 4268 CCCCTTCCCCTTCCCGCAGCTTACGCTTCTTCCAGGGGAGTACTCAGGACTCAGCT 4327
Db 299 CCCCTTCCCCTTCCCGCAGCTTACGCTTCTTCCAGGGGAGTACTCAGGACTCAGCT 240
QY 4328 AGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCCTCAACCTCTCTGGGGCAG 4387
Db 239 AGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCCTCAACCTCTCTGGGGCAG 180
QY 4388 GGGAGCGCGAGACTCCCGTGGGAGAAGCTCATTTCCCAATCTTGCAGAGAGCTTTTGTG 4447
Db 179 GGGAGCGCGAGACTCCCGTGGGAGAAGCTCATTTCCCAATCTTGCAGAGAGCTTTTGTG 120
QY 4448 CAGTGTCCACATTTGAGTCAGACTGCTCCCGGGGAGAGAGCCCGGGCCCCCAGGACATAA 4507
Db 119 CAGTGTCCACATTTGAGTCAGACTGCTCCCGGGGAGAGAGMCCCGGGCCCCCAGGACATAA 60
QY 4508 AGAATCGCAGCTTGGTACTCAGAGTCTGGTGTGTAGAGAACTCTTTTGTAAAGCAATAA 4567
Db 59 AGNACTGCGCTTGGNACTGCAGANN--NGNTTNNAGAGAAGTCTTAAAGNNAARAA 1
```

RESULT 8  
AL556130

LOCUS AL556130 1133 bp mRNA linear EST 30-MAR-2004  
DEFINITION Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DK011Y14 5-PRIME, mRNA sequence.

ACCESSION AL556130  
VERSION AL556130.3 GI:45860849  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1. (bases 1 to 1133)  
Li, W.B., Gruber, C., Jessee, J. and Pollayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:31277934.

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
8426.r

For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DK011Y14&q=8426.r.

FEATURES  
Location/Qualifiers  
1..1133  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK011Y14"  
/cell\_type="HELA CELLS COT 25-NORMALIZED"  
/cell\_line="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6  
vector. Library was normalized."

ORIGIN  
Query Match 19.9%; Score 916; DB 1; Length 1133;  
Best Local Similarity 95.3%; Pred. No. 1.9e-219;  
Matches 1058; Conservative 4; Mismatches 34; Indels 14; Gaps 11;

Qy 2922 AGCTTCAGCTTGTCACCAAGAGCGGCGCCCTGTCGCTACGACATGGAGTGCAGCA 2981  
Db 1 AGCCTTCAGCTTGTCACCAAGAGCGGCGCCCTGTCGCTACGACATGGAGTGCAGCA 60  
Qy 2982 CTCGGGATCGACCTGCAGTCACACTGGCCCTGATGGCAGCTTCGGCTGGAGCTGGAG 3041  
Db 61 CTCGGGATCGACCTGCAGTCACACTGGCCCTGATGGCAGCTTCGGCTGGAGCTGGAG 120  
Qy 3042 GGTCAAGCATGGCAGCTGGAGAAACAGGCCCTAACCTTGCCTCCACCGCGGTCCACA 3101  
Db 121 GGTCAAGCATGGCAGCTGGAGAAACAGGCCCTAACCTTGCCTCCACCGCGGTCCACA 180  
Qy 3102 CTGCGGAAAGCAGCTTCCTCTCGGTGACAGATGCTGCCCTGAACACAGAGCTCAGC 3161  
Db 181 CTGCGGAAAGCAGCTTCCTCTCGGTGACAGATGCTGCCCTGAACACAGAGCTCAGC 238  
Qy 3162 CGTTCCAGGGGATTCGACAGCCCGCCCTGATGGGAAACAGGGCTCGCAGCAGCA 3221  
Db 239 CGTTCCAGGGGATTCGACAGCCCGCCCTGATGGGAAACAGGGCTCGCAGCAGCA 297  
Qy 3222 AGGTGGGGGAGAGAGATGCTCCAGGATTTTACACCTGAGCCCTGCCCCACCTGCT 3281  
Db 298 AGGTGGGGGAGAGAGATGCTCCAGGATTTTACACCTGAGCCCTGCCCCACCTGCT 357  
Qy 3282 GAAAAAACA-TCCGCGACGTGAAGAGACAGAGAGAGATGCGAGGTTACTCGGGAAA 3340  
Db 358 GAAAAAACA-TCCGCGACGTGAAGAGACAGAGAGAGATGCGAGGTTACTCGGGAAA 416

Qy 3341 CAAAACAGGGATCTTTTCTGCCCCCTGCTCAGTCGAGTTGGCCTGACCCCGTTGGATCA 3400  
Db 417 CAAAACAGG----ATCTTCTCTGCCCCCTGCTCAGTCGAGTTGGCCTGACCCCGTTGGATCA 472  
Qy 3401 GTGACCACTTTGTTGGCAGACAGGGGAGAGCAGCTTCCAGCCCTGGGTACAGAGGGGTGGGC 3460  
Db 473 GTGACCACTTTGTTGGCAGACAGGGGAGAGCAGCTTCCAGCCCTGGGTACAGAGGGGTGGGC 532  
Qy 3461 GAGCCCTTCGCCCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCA 3520  
Db 533 GAGCCCTTCGCCCCCTCACCCCTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCA 592  
Qy 3521 AACTCAGGTTCAAGTCAGAGAACCAAGGTGATGCCCCCGTATGAGTAAAGGGGGCC 3580  
Db 593 AACTCAGGTTCAAGTCAGAGAACCAAGGTGATGCCCCCGTATGAGTAAAGGGGGCC 651  
Qy 3581 CTCTAAACCCCTTGGCTGGCTCACCCTGGCCAGCTCACCCTTTTGGGTGAGGGGAAA 3640  
Db 652 CTCTAAACCCCTTGGCTGGCTCACCCTGGCCAGCTCACCCTTTTGGGTGAGGGGAAA 711  
Qy 3641 GAATGCTGACCCCTGGGAAAGGCTCCCT-GGTAGATATACACACACTTTTCAAGTTGTTCG 3699  
Db 712 GAATGCTGACCCCTGGGAAAGGCTCCCTGGGTAGATATACACACACTTTTCAAGTTGTTCG 771  
Qy 3700 AACACAGGTCCTGAGTTGACCTCTGTTTCAGCCCAAGGACCAAGAGGTGTGAAGTCAA 3759  
Db 772 AACACAGGTCCTGAGTTGACCTCTGTTTCAGCCCAAGGACCAAGAGGTGTGAAGTCAA 831  
Qy 3760 GTGGTTCTCAGTCCCGACAGATGTGCCCCCTTGGTGTGCTGGTACACACTTCTCCCGAGAGC 3819  
Db 832 GTGGTTCTCAGTCCCGACAGATGTGCCCCCTTGGTGTGCTGGTACACACTTCTCCCGAGAGC 891  
Qy 3820 AGCAGGCCCCGAGCCCTTCAGGCCCCAGACTGCTGCCCAGACTCCTGCGACTCAGTTCC 3879  
Db 892 AGCAGGCCCCGAGCCCTTCAGGCCCCAGACTGCTGCCCAGACTCCTGCGACTCAGTTCC 951  
Qy 3880 TCATCTCTAAAGGTG-AAGGGTGTATGACGATATGCTCCAGCAAGACAGTCTGTGGATGG 3938  
Db 952 TCATCTCTAAAGGTGAAAGGGTGTATGACGATATGCTCCAGCAAGACAGTCTGTGGATGG 1011  
Qy 3939 ACATGATCAGTCTAAAGAAAGCAGCAGAGAGAGAGCTCGCGGCCCGCCAGCCCACTATC 3998  
Db 1012 ACATGATCAGTCTAAAGAAAGCAGCAGAGAGAGAGCTCGCGGCCCGCCAGCCCACTATC 1070  
Qy 3999 AGTCTCCAGGCTGCTGTTCCCGAGAGCAC 4028  
Db 1071 AGTCT-CAGAGTCTGTTCCAGAAAGACAS 1099

RESULT 9  
LOCUS BM458601  
DEFINITION BM458601 1140 bp mRNA linear EST 05-FEB-2002  
AGENCOURT\_6413524 NIH\_MGC\_85 Homo sapiens cdna clone IMAGE:5497658  
5', mRNA sequence.  
ACCESSION BM458601  
VERSION BM458601.1 GI:18507641  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1. (bases 1 to 1140)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-x@mail.nih.gov  
Tissue Procurement: Lou Straudt  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be



found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12128 row: c column: 03  
High quality sequence stop: 724.

FEATURES

source  
1. .1140  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5497658"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 85"  
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.867 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 19.1%; Score 876.2; DB 4; Length 1140;  
Best Local Similarity 94.1%; Pred. No. 2.1e-209; Mismatches 0; Indels 15; Gaps 11;  
Matches 1030; Conservative 0;  
QY 2556 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGACAAATTC 2615  
DB 44 AGAATTATTCCTCAACAGCCTGTCCAGCCATTTTCTCTGGAGGAGCAGGACAAATTC 103  
QY 2616 CTCGTGCTCAGCATCGACAGCCTCTCTCTGCGATGACATGAGAGAACCCATCAAA 2675  
DB 104 CTCGTGCTCAGCATCGACAGCCTCTCTCTGCGATGACATGAGAGAACCCATCAAA 163  
QY 2676 GGCCTCTCAAGCTCGGGGACACCTGAGCTCAGCGTACACTCTCTGGAGCAGCAGGC 2735  
DB 164 GGCCTCTCAAGCTCGGGGACACCTGAGCTCAGCGTACACTCTCTGGAGCAGCAGGC 223  
QY 2736 CGAGGCTCGAAGCTCAGCTGGAACATGCTGTGCGCGCGCGGCCACCCAGCACCC 2795  
DB 224 CGAGGCTCGAAGCTCAGCTGGAACATGCTGTGCGCGCGCGGCCACCCAGCACCC 283  
QY 2796 AAGCTATTTCAATGTGTGAAGTCCAATACAGTCTCTTAATGTGTAACACCTGCAT 2855  
DB 284 AAGCTATTTCAATGTGTGAAGTCCAATACAGTCTCTTAATGTGTAACACCTGCAT 343  
QY 2856 CGGAGGTTCCACCGGTCAGAGTGGAGACATCCCACTGCGATCAGCAGCCAGATCCC 2915  
DB 344 CGGAGGTTCCACCGGTCAGAGTGGAGACATCCCACTGCGATCAGCAGCCAGATCCC 403  
QY 2916 AGCTGAGCTTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGATGGAGT 2975  
DB 404 AGCTGAGCTTTCAGCTTGGTCAACAAAGACGGGAGCCTGTTCTGCTACGATGGAGT 463  
QY 2976 GCCAGACTGGGCATCGACTCGATGACATCGGCCCTGTATGAGAGTTGCGCTTGAG 3035  
DB 464 GCCAGACTGGGCATCGACTCGATGACATCGGCCCTGTATGAGAGTTGCGCTTGAG 523  
QY 3036 CTGGAGGTTCAAGCATGGCCAGTGGAGAACAGGCGCTTAACCTGCCCTCCACCGCGGC 3095  
DB 524 CTGGAGGTTCAAGCATGGCCAGTGGAGAACAGGCGCTTAACCTGCCCTCCACCGCGGC 583  
QY 3096 TCCACACTGCCGGAAGCAGCCTTCTCTGCTCGGTGACGATGCTGCCCTCGAAACACAGG 3155  
DB 584 TCCACACTGCCGGAAGCAGCCTTCTCTGCTCGGTGACGATGCTGCCCTCGAAACACAGG 641  
QY 3156 CTCAGCGTTTCCAGGGATTCGCCAGCCCCCGCTTCACAGTGGGAAACAGGCGCTCGCA 3215  
DB 642 CTCAGCGTTTCCAGGGATTCGCCAGCCCCCGCTTCACAGTGGGAAACAGGCGCTCGCA 700  
QY 3216 CGAGCAAGTGGGGGACAGCAATGCTCCAGATTTTCAACCTGAGCCTGCGCTGCCAC 3275  
DB 701 CGAGCAAGTGGGGGACAGCAATGCTCCAGATTTTCAACCTGAGCCTGCGCTGCCAC 760  
QY 3276 CCTGCTGAAAAAACA-TCCGCCACGTTGAAGACAGAGGAGGATGGCAGGATTTACCTG 3334

DB 761 CCTGCTGAGAAAACACTCCGCCACGTAAGAGACAG-AGGAGATGGCAGGAGTTACCTC 819  
QY 3335 GGGAAACAAAACAGGATCTTTTCTGCCCTCTCCAGTCGAGTCGAGTCGACCGCTT 3394  
DB 820 GGGGAAACAAAACAGG---ATCTTCTCTGCCCTCGCTCAGTCAGTTGGCCTTACCCGCTT 876  
QY 3395 GGATCAGTGACCATTTGTTGGCAGACAGGAGAGCAGCTTCCAGCCT-GGGTCAGAAAG 3453  
DB 877 GGATCAGTGCCCATTTGCTGGCAGACA-GGAGAGCAGCTTCCAGCTGGGGTCAAAGG 935  
QY 3454 GGTGGGCGAG-CCCTTCGGGCCCTCACCTCCAGGCTCTGTGAGAGTGTCAAGTGTGTA 3512  
DB 936 GGTGGGCGAGCCCTTGGGCCCTCACCTCCAGGCTGCTGTGAAAGTNTCAAGTGTGTA 995  
QY 3513 AGGGCCCAAACTCAGTTTCAGTCGAGAACACAGGTCAGCAGGTATGCC--CGGCCCTAGGT 3570  
DB 996 AGGGCCCAAACTCAGTTTCAGTCGAGAACACAGGTCAGCAGGTATTCGCCGCCCTTAGGT 1055  
QY 3571 TAAAGGGGCGCTCTAAACCCCTTGCTGCTGCTCACCTGCGCCAGCT--CACCCCTTTTGGG 3628  
DB 1056 TAGGGGGGCGCTCTAAACCCCTTGCTGCTGCTTAACGGNCAGGTTACCCCTTTTGGG 1115  
QY 3629 TGTAGGGGAAAGRA 3643  
DB 1116 GGAAGGGGAAAGAA 1130  
RESULT 10  
BU541005  
LOCUS  
DEFINITION BU541005 955 bp mRNA linear EST 13-SEP-2002  
AGENCY: NIH\_MGC 42 Homo sapiens cDNA clone IMAGE:6572558  
5' mRNA sequence.  
ACCESSION BU541005  
VERSION BU541005.1 GI:22851446  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 955)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LCM2764 row: n column: 14  
High quality sequence start: 31  
High quality sequence stop: 630.  
Location/Qualifiers  
1. .955  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6572558"  
/tissue\_type="epithelioid carcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 42"  
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;  
Site 2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).



Note: this is a NIH\_MGC Library. | "

ORIGIN	BM454415	LOCUS	DEFINITION	BM454415	1038 bp	mrna	linear	EST 05-FEB-2002
			AGENCOURT 6402274 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5497584					
			5', mRNA sequence.					
		ACCESSION	BM454415					
		VERSION	BM454415.1	GI:18503455				
		KEYWORDS	EST.					
		SOURCE	Homo sapiens (human)					
		ORGANISM	Homo sapiens					
			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
		REFERENCE	1 (bases 1 to 1038)					
		AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.					
		TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
		JOURNAL	Unpublished (1999)					
		COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaaps-r@mail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12127 row: p column: 01 High quality sequence stop: 595.					
		FEATURES	source					
			1..1038					
			/organism="Homo sapiens"					
			/mol_type="mRNA"					
			/db_xref="taxon:9606"					
			/clone="IMAGE:5497584"					
			/tissue_type="lymphoma, cell line"					
			/lab_host="DH10B (phage-resistant)"					
			/clone_lib="NIH_MGC_85"					
			/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."					
		ORIGIN						
			Query Match	17.6%	Score	808.8	DB 4	Length 1038
			Best Local Similarity	96.2%	Pred. No.	2.1e-192		
			Matches	883	Conservative	0	Mismatches	27
							Indels	8
							Gaps	5
Qy	630	TTGGAAGGAAAGCGTCCGAGCAAGCCCGAAGAAACGGAAGAAAGAGAGCTCAAGTC	689					
Db	2	TTGGAAGGAAAGCGTCCGAGCAAGCCCGAAGAAACGGAAGAAAGAGAGCTCAAGTC	61					
Qy	690	CCTGGCTCATGAGGAGTGGCTTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAG	749					
Db	62	CCTGGCTCATGAGGAGTGGCTTGGCCAAACCCCTCCAGGACCCCTGAGCAGGAGAG	121					
Qy	750	CTGCACCATCCAGTCCGAGGAGTGTCTCCACTCGGCGCCCATATGTTAGAAACAC	809					
Db	122	CTGCACCATCCAGTCCGAGGAGTGTCTCCACTCGGCGCCCATATGTTAGAAACAC	181					
Qy	810	CCGCGAGTTCACCAAGCCTCTGAAGAAACAGGGCTTGGGCAACTCTGTTTAAAGCAGCT	869					
Db	182	CCGCGAGTTCACCAAGCCTCTGAAGAAACAGGGCTTGGGCAACTCTGTTTAAAGCAGCT	241					
Qy	870	TGGCGAGGGCTACGGCGGCTCTGCTCGATCAGAACTCCACAACTGATCAGCCCTT	929					
Db	242	TGGCGAGGGCTACGGCGGCTCTGCTCGATCAGAACTCCACAACTGATCAGCCCTT	301					
Qy	930	GCAATGTCTGAACCAACGCTGTGGAACTGCACACCCCGAGGAGGCGCCCTTGCCTT	989					
Db	302	GCAATGTCTGAACCAACGCTGTGGAACTGCACACCCCGAGGAGGCGCCCTTGCCTT	361					
Qy	990	GCCACGACCCCTTCCCTTATAGCAGACTGCTCCTTCCCTTCCCTTCCCTTCCCTTCCCA	1049					
Db	362	GCCACGACCCCTTCCCTTATAGCAGACTGCTCCTCATCCCTTCCCTTCCCTTCCCTTCCCA	421					

RESULT 11

QY 1050 GCCTGGAACCTCACCTCTTGAGTCCCTTCTGGGCAAACTGGCCTGTGTAGACAGCCA 1109  
Db 422 GCCTGGAACCTCACCTCTTGAGTCCCTTCTGGGCAAACTGGCCTGTGTAGACAGCCA 481  
QY 1110 GAAACCTTTGCCCTGACCCACACCTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCT 1169  
Db 482 GAAACCTTTGCCCTGACCCACACCTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCT 541  
QY 1170 GCCTGGCCACACCTTGAGCCACAGTGCCTGTCTGTGTGGCCCATGAGAACTTTTCTGT 1229  
Db 542 GCCTGGCCACACCTTGAGCCACAGTGCCTGTCTGTGTGGCCCATGAGAACTTTTCTGT 601  
QY 1230 GGAGGAATACCTAGTGCATGTCTGCAAGGACGCTGAGCTCAAGCCAGGCCACAGCCT 1289  
Db 602 GGAGGNATACCTAGTGCATGTCTGCAAGGACGCTGAGCTCAAGCCAGGCCACAGCCT 661  
QY 1290 GACAGCCTGGCCAAAGACTGGGACGACACGGGGCTCCAGATCCCGGAGCCAGCCCAA 1349  
Db 662 GACAGCCTGGCCAAAGACTGGGACGACACGGGGCTCCAGATCCCGGAGCCAGCCCAA 721  
QY 1350 RACTGAGGACAGCGGG-TGTCCTGCTCACTGAGAACTCAAGCCAG-TGATTTATGAG 1407  
Db 722 RACTGAGGACAGCGGGTGCCCTGCTCACTGAGAACTCAAGCCAGTTGATTTATGAG 781  
QY 1408 TACCAGAGAAAGTCCACTGGGCCACGACACGCTCCGCTCGGCGAGAGGCT-CCTTCGG 1466  
Db 782 TACCAGAAAGTCCACTGGGCCACGACACGCTCCGCTCGGCGAGAGGCTCCCTTCGG 841  
QY 1467 AGAGGTGCAC--AGGATGAGGACAGCAGACTGGC---TTCAGTGGCTGTCAAAAAG 1521  
Db 842 AGAGGGCACACAGGATGGGAGACAGCAGACTGGCCTTCAGTGGCCTGTCCAAAAG 901  
QY 1522 GTGCCCTGAGATTT 1539  
Db 902 GGGCGCTGGAAGAATT 919

RESULT 12  
BX331085/c  
LOCUS  
DEFINITION BX331085 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DK011YM10 3-PRIME, mRNA sequence.  
ACCESSION BX331085  
VERSION BX331085.2 GI:46263202  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 909)  
Li W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On May 1, 2003 this sequence version replaced gi:30307969.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.

FEATURES  
source  
1..909  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DK011YM10"  
/cell\_type="HELA CELLS COT 25-NORMALIZED"  
/cell\_line="HELA"  
/clone\_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 17.1%; Score 787.4; DB 5; Length 909;  
Best Local Similarity 95.7%; Pred. No. 5.1e-187; Indels 6; Gaps 6;  
Matches 872; Conservative 0; Mismatches 33;

QY 3679 CCACACTTTTTCAGGTTGTGCAACACA-GGTCCTGAGTTG-ACCTCTGTTTCAGCAAGG 3736  
Db 909 CCAGGCTTTTCAGGTTGTGCAACANACAGGGTCTGAGTTGAACCTCTGTCAGCCAAAGG 850  
QY 3737 ACCAAGAAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCC-AGACATGTGCCCTTTGTCTG 3795  
Db 849 GCCCCAAAAAAGGTGTGTAATGAAATGGGTTTTCAGTCCCCAAGACATGGGCCCTTTGGCTG 790  
QY 3796 CTGGGCTACACTCTTTCCTCCAGACAGCAGGCCCCGAGGCCCTTTCAGGGCCAGCACTGGCC 3855  
Db 789 CGGGCTACCACTCTTTCCTCCAGACAGCAGGCCCCGAG-CCCTTCAGGCCAGCACTGGCT 731  
QY 3856 CAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAAGGATATGCC 3915  
Db 730 CAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAAGGATATGCC 671  
QY 3916 TGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAGCAGAGAGAGC 3975  
Db 670 TGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAGCAGAGAGAGC 611  
QY 3976 -TCGGGCCCCCGAGCCCACTATCAGTGTCCAGCGTGTGTTCCCGCAGAGCAGAGTCA 4034  
Db 610 CTCGGGCCCCCGAGCCCACTATCAGTGTCCAGCGTGTGTTCCCGCAGAGCAGAGTCA 551  
QY 4035 GCATCACACTGACACTCACCTGCGCTCCCTGCGCCAGAGGGTACTGCCAGCGCACTT 4094  
Db 550 GCATCACACTGACACTCACCTGCGCTCCCTGCGCCAGAGGGTACTGCCAGCGCACTT 491  
QY 4095 TGCACCTGTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGCGAGGCGCA 4154  
Db 490 TGCACCTGTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGCGAGGCGCA 431  
QY 4155 GTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACAGTCTTTGAGCTG 4214  
Db 430 GTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACAGTCTTTGAGCTG 371  
QY 4215 TCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGATAGCACCCCTTC 4274  
Db 370 TCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGATAGCACCCCTTC 311  
QY 4275 CCCATTGCCCGAGCTTAGCCTCTTCCAGGGGAGTACTCAGGACTCAGTAGCATTA 4334  
Db 310 CCCATTGCCCGAGCTTAGCCTCTTCCAGGGGAGTACTCAGGACTCAGTAGCATTA 251  
QY 4335 AATCAGCTGTGAATCGTTCAGGGGGTGTCTGTAGCTCAAACTCTCTGGGGCAGGGGAGCG 4394  
Db 250 AATCAGCTGTGAATCGTTCAGGGGGTGTCTGTAGCTCAAACTCTCTGGGGCAGGGGAGCG 191  
QY 4395 CGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCGCAAGACAGCCTTTTGTCCAGTGT 4454  
Db 190 CGAGACTCCGTGGGAGAGCTCATTTCCACATCTTTCGCAAGACAGCCTTTTGTCCAGTGT 131  
QY 4455 CCACATTGAGTCAGACTGTCTCCGGGGAGAGAGCCCGGCCCCAGCAGCATTAAGAACTG 4514  
Db 130 CCACATTGAGTCAGACTGTCTCCGGGGAGAGAGCCCGGCCCCAGCAGCATTAAGAACTG 71  
QY 4515 CAGCCTTGGTACTGCAGAGTCTGGGTGTAGAGAACTTTTGTAAAGCAATAAAGTTTGGG 4574  
Db 70 CAGCCTTGGTACTGCAGAGTCTGGGTGTAGAGAACTC-TTGTAAAGCAATAAAGTTTGGG 12  
QY 4575 GTGATGACAA 4585  
Db 11 GTGATGACAA 1

```
RESULT 13
BG775354/c
LOCUS      602649808T1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:4761048 3',
DEFINITION mRNA sequence.
ACCESSION  BG775354
VERSION     BG775354.1 GI:14045671
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: DCTD/BTP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LNCM1612 row: o column: 01
            High quality sequence start: 2
            High quality sequence stop: 864.
            Location/Qualifiers
                1..1037
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4761048"
                /tissue_type="carcinoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 40"
                /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
                Site 2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Library constructed by
                Ling Hong in the laboratory of Gerald M. Rubin (University
                of California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."

ORIGIN
Query Match      16.5%; Score 756.2; DB 4; Length 1037;
Best Local Similarity 93.3%; Pred. No. 4e-179;
Matches 889; Conservative 0; Mismatches 53; Indels 11; Gaps 9;

Qy 3588 CCCCTTGCTGGCTCCTGCGGCTCAGCTGCGGCTCAGCTCCCTTTTGGGTGTAGGGGAAAAGATGCC 3647
Db 949 CCTCTAACCCCTGCTGGCCACACAGCGGCTACCCCTTTTGGTCTAGGGAACATGAACGCC 890

Qy 3648 TGACCCCTGGGAGGCTCCCTGGTAGATACACCACTTTTCAGGTGTTGCAACACAGG 3707
Db 889 TGAACCCCGGGAAG--CTCCCTGTGTAGATACACCACTCTTCA--GTTGTTGCAACACAGG 833

Qy 3708 TCCTGAGTTGACCTCTGTGTTCA--GCCAAGGACCAAGAGGTGTAGTGAAGTGGTTC 3766
Db 832 TCTGAGTTGAGTCTCCCGTTTCAGCCAGGACCAAGAGGTGTGTAA--TGAATGTTTC 775

Qy 3767 TCAGTCCCAACATATG--CCCTTTGCTGT--GGCTACCACTCTTCCCCAGACAGCAG 3824
Db 774 TCAGTCCCAACAGGCGCCCTTGTGCTGGGCTACCACTCTTCCCCAGACAGCAG 715

Qy 3825 GCGCGAGCCCTTCAGGCGCCAGCACTGCCCGACACTCGCTGGCACTCAGTTCCTCATC 3884
Db 714 GCGCGAGCCCTTCAGGCGCCAGCACTGCCCGACACTCGCTGGCACTCAGTTCCTCATC 655
```

```
Qy 3885 TGTAAAGGTGAAGGGTGATGTCAGGATATGCTGTGACAGAAACAGTCTGTGGATGACATGA 3944
Db 654 TGTAAAGGTGAAGGGTGATGTCAGGATATGCTGTGACAGAAACAGTCTGTGGATGACATGA 595

Qy 3945 TCAGTGTCTAAGGAAGCAGCAG--AGAGAGAG--TCCGCGCGCCAGCCCACTATCAGTG 4002
Db 594 TCAGTGTCTAAGGAAGCAGCAGTAGAGAGCGTCCGCGCGCCAGCCCACTATCAGTG 535

Qy 4003 TCCACGCTGTGTTGTTCCCGCAGACACAGCTCAGCATCACACTGACCTCACCCTGCCCTG 4062
Db 534 TCCACGCTGTGTTGTTCCCGCAGACACAGCTCAGCATCACACTGACCTCACCCTGCCCTG 475

Qy 4063 CCCCTGGCCAGAGGTTACTGCCGACGGCAGCTTTTGCACCTCTGATGACCTCAAAGCACTTTC 4122
Db 474 CCCCTGGCCAGAGGTTACTGCCGACGGCAGCTTTTGCACCTCTGATGACCTCAAAGCACTTTC 415

Qy 4123 ATGCTCCCTCTCGCAGGCGCAGGCGCAGTGACACTGTAGGAGCATAGCAAGCCAG 4182
Db 414 ATGCTCCCTCTCGCAGGCGCAGGCGCAGTGACACTGTAGGAGCATAGCAAGCCAG 355

Qy 4183 GAGATGGGTTGAAGGACACAGCTTTGAGCTGTCCACATGATGTGACTCCTCAAACCTC 4242
Db 354 GAGATGGGTTGAAGGACACAGCTTTGAGCTGTCCACATGATGTGACTCCTCAAACCTC 295

Qy 4243 TTCAGATTTCTTAAAGATAGCACCCCTTCCCATTTGCCCACTGAGCTCTTCTCC 4302
Db 294 TTCAGATTTCTTAAAGATAGCACCCCTTCCCATTTGCCCACTGAGCTCTTCTCC 235

Qy 4303 CAGGGAGCTACTCAGGACTCAGTAGCATTAATGACTGTGAATCGTCAGGGGGTGTTC 4362
Db 234 CAGGGAGCTACTCAGGACTCAGTAGCATTAATGACTGTGAATCGTCAGGGGGTGTTC 175

Qy 4363 TGCTAGCTCAACCTCTCTGGGCGAGGGGAGCCCGAGACTCCGTTGGGAGAACTCATTTCCC 4422
Db 174 TGCTAGCTCAACCTCTCTGGGCGAGGGGAGCCCGAGACTCCGTTGGGAGAACTCATTTCCC 115

Qy 4423 ACATTTTCCCAAGACACCTTTTGTCCAGCTGTCCACATTTGAGTCAGAGTGCTC--CCGGG 4481
Db 114 ACATTTTCCCAAGACACCTTTTGTCCAGCTGTCCACATTTGAGTCAGAGTGCTC--CCGGG 55

Qy 4482 AGAGAGCCCGCGCCCGCCAGCACATAAAGAACTGCGCTTGGTACTCGAGGT 4534
Db 54 AGAGAGCCCGCGCCCGCCAGCACATAAAGAACTGCGAGACTTGGTACTCGAGGT 2

RESULT 14
BG822015
LOCUS      602726122F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4865789 5',
DEFINITION mRNA sequence.
ACCESSION  BG822015
VERSION     BG822015.1 GI:14169602
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: NIH Intramural Sequencing Center
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LNCM1731 row: k column: 06
            High quality sequence stop: 748.
            Location/Qualifiers

FEATURES
```



Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: <http://genome.sgc.riken.jp/>  
URL: <http://fantom.sgc.riken.jp/>  
Location/Qualifiers  
source  
1. 2964  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:A230066D12"  
/db\_xref="taxon:10090"  
/clone="A230066D12"  
/sex="male"  
/tissue\_type="hypothalamus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
misc\_feature  
1. 2964  
/note="mitogen-activated protein kinase kinase 14  
(MGP|MGI:1858204, GB|NM\_016896, evidence: BLASTN, 100%,  
match=2829)"  
ORIGIN  
Query Match 16.1%; Score 741.6; DB 3; Length 2964;  
Best Local Similarity 82.6%; Pred. No. 2.7e-175;  
Matches 886; Conservative 0; Mismatches 179; Indels 7; Gaps 3;  
Qy 329 GGAAGAAACAGAGTCCGCTCTACAGCTTGAGCGGTGAGGAGAGAGCCGTGTGTCGCG 388  
Db 1 GTAAGAAGCAGAGCTGCATCTTCAAGCTTGAGCGGTGAGGAGAGAGCCGTGTGTCG 60  
Qy 389 GAAAGTGGAGATCTGATGACGTGATTACCAAGGGCAGAGCCCAAGAGGCTCCGAGG 448  
Db 61 GGAAGTGGAGATCTTAAACGACGTGATCACCAAGGCGACAGCCCAAGGACGGCTCTGAGG 120  
Qy 449 CAGGCGCAGCTGCGCATCTCTATCATCGCCAGGCTGAGTGTGAGAAATAGCCAGAGTTCA 508  
Db 121 GAGAACCAACCGGCGCATCTCCATCATCGCCAGGCTGATGTGAGAAATAGCCAGAGTTCA 180  
Qy 509 GCGCCACCTTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGA 568  
Db 181 GCCCACCTCTCAGAGCGCATTTTCATCGCGGGTACACAGCAGTACAGCCAGTCTGAGA 240  
Qy 569 GTCTTGATCAGATCCCAACAAATGTGGCCCATGTCTACAGAGGGCAAAATGGCCCGTGTGT 628  
Db 241 GTCTCGATCAAAATCCCAACAAATGTGGCCCATGTCAACTGAAAGGCAAAATGGCCCGTGT 300  
Qy 629 GTTGGAGGGAAAGCTGCGCAGCAAGAGCCGGAAGAAACGGAAGAAAGAGCTCAAAAGT 688  
Db 301 GCCGGAGGGGAAAGAGTCAAGCAAGAGCCGGAAGAAACGTAAGAAAGAGGTGGAAGT 360  
Qy 689 CCCTGGGTTCATCAGGAGTGGCTTTGGCCAAACCCCTCCAGGAGCCCTCAGCAGGAGA 748  
Db 361 CACTGGCCACAGGAGGTGGCTTTAGCCAAAGCCCTGCCCAGAAACCCCTGAGCAAGAGA 420  
Qy 749 GCTGACCATCCAGTGCAGAGGATGAGTCTCACTCGCGGCCCCATATGTTTAAAGACA 808  
Db 421 GCTGTACCATCCAGTACAGGAAGATGAGTCTCACTAGGCAACCTCTATGCCAGAAATG 480  
Qy 809 CCCCGAGTTCACCAAGCCTCTGAAGGAACAGAGCCCTTGGCAACTCTGTTTAAAGCAGC 868  
Db 481 TCTCCAGTTCACCAAGCCTCTGCGGGGACCCAGGCTTGGCCACCTGTGCTTTAAGAAAC 540  
Qy 869 TTGCGAGGGCTACGGCGGCTCTGCTCGATCAGAACTCCAAACCTGATCAGCCCT 928  
Db 541 AAGATGAAGGCTCGGACCGGTACTGCTCGACCAAGAACTCCAAACTGATCAGCCCT 600  
Qy 929 TGCAATGTGCAACCAAGTGTGGAACCTGCAACCCAGGACCGGAGGCCCTCGCCCC 988  
Db 600

601 TGCATGTCTTAAACCAGCTGTGGAAACTCCACACCCCGAGGCCACAGGCCCCCGGCCCC 660  
989 TGCCC-ACGACACCCCTTCCCTATAGCAGAGTGCCTCATCCCTTCCATTCCACCCCTCTC 1047  
661 CACCGGACTCACCCCTTCCCTTACAGCGGAATGCCCATCTCTTCCCATTTACCCCTG 720  
1048 CAGCCCTGGAAACCTCACCCCTCTGGAGT---CCTTCTGGGCAAACTGGCTGTGTAGAC 1104  
721 GAGCCCTGGAAACCTTATATGCTGGACTCTGCCGTCTCTGGACAAACTAGCCGGTGTGAGC 780  
1105 AGCCAGAAACCTTTCCTGACCCCA---CACCTGAGCAAACTGGGCTGTGTAGACAGTCCA 1161  
781 GGCAGCGGCTCTGCTGCTGCCACCCGATCTAAGCCAACTGGCCCATGGAGACAGTCAAG 840  
1162 AAGCCCTGCTGCTGCCCCACACCTGGAGCCAGCTGCTGTCTGTGTGTCCTATGAGAG 1221  
841 AAGCCGCTGCTGCTGCCCCACACCTGGAGTCCAGCTGCCGTCTCGGGGTGCCCTAGAAAAG 900  
1222 TTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGACAGCTGAGCTCAAGCCAGGCC 1281  
901 GTTCCCGTGGAGGAATACCTGGTGCATGCGCTCCAAGGAAGTGTGAGCTCAGGCCAGGCC 960  
1282 CACAGCTGACCCAGCTGGCCAAAGACCTGGGCGAGACAGCGGCTCCAGATCCCGGGAGGCC 1341  
961 CACAGCTGGCCAGCTGGCTAAGACATGTCTCGGGAAGCGCCAAAGCTCAGAGGCTC 1020  
1342 AGCCCCAAACTGAGGACAAACGAGGGTGTCTCTGCTCACTGAGAAACTCAAGC 1393  
1021 GGCCCCGAAACTGAGGACAAACGAGGGGTCTCTGCTTACTGAGGTAACTAACC 1072

Search completed: June 10, 2005, 20:27:55  
Job time : 9696.08 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 14:23:14 ; Search time 1746.62 Seconds  
(without alignments)  
16311.995 Million cell updates/sec

Title: US-09-155-676B-6  
Perfect score: 4596  
Sequence: 1 agcgggggactgtgccgtg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
  - 2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
  - 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
  - 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*
  - 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*
  - 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq.\*
  - 19: /cgn2\_6/ptodata/2/pubpna/US10G\_PUBCOMB.seq.\*
  - 20: /cgn2\_6/ptodata/2/pubpna/US10H\_PUBCOMB.seq.\*
  - 21: /cgn2\_6/ptodata/2/pubpna/US10I\_PUBCOMB.seq.\*
  - 22: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
  - 23: /cgn2\_6/ptodata/2/pubpna/US11A\_PUBCOMB.seq.\*
  - 24: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*
  - 25: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
  - 26: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4578.4	99.6	4596	10	US-09-981-397A-17
2	4578.4	99.6	4596	17	Sequence 17, Appl
3	4578.4	99.6	4596	18	Sequence 975, App
4	4578.4	99.6	4596	19	Sequence 975, App
5	4440	96.6	4684	13	Sequence 294, App
6	2837.6	61.7	2844	9	Sequence 887, App
7	2042.6	44.4	2828	13	Sequence 2, Appli
					Sequence 884, App

8	1556.2	33.9	73967	13	US-10-087-192-886	Sequence 886, App
9	1304	28.4	1385	18	US-10-296-115-97	Sequence 97, Appl
10	480.2	10.4	42566	13	US-10-087-192-883	Sequence 883, App
c 11	349.4	7.6	2475	17	US-10-108-260A-1754	Sequence 1754, App
12	291.6	6.3	6673	19	US-10-433-793-141	Sequence 141, App
13	267.4	5.8	499	10	US-09-918-995-2585	Sequence 2585, Ap
c 14	243.6	5.3	6673	19	US-10-433-793-142	Sequence 142, App
15	220.6	4.8	296	9	US-09-783-590-11852	Sequence 11852, A
16	171.8	3.7	500	16	US-10-029-386-8637	Sequence 8637, Ap
17	166.8	3.6	293	16	US-10-029-386-22365	Sequence 22365, A
c 18	88.4	1.9	2348	18	US-10-380-127A-3	Sequence 3, Appli
c 19	88.4	1.9	455	20	US-10-357-930-22052	Sequence 22052, A
c 20	88.4	1.9	455	20	US-10-357-930-27909	Sequence 27909, A
c 21	88.4	1.9	476	20	US-10-357-930-12325	Sequence 12325, A
c 22	88.4	1.9	504	20	US-10-357-930-33496	Sequence 33496, A
c 23	88.4	1.9	504	20	US-10-357-930-42420	Sequence 42420, A
c 24	83.8	1.8	3332	13	US-10-000-864-11	Sequence 11, Appl
25	82.2	1.8	1935	13	US-10-000-864-5	Sequence 5, Appli
c 26	73.8	1.6	463	20	US-10-357-930-3156	Sequence 3156, Ap
27	66.2	1.4	1300	18	US-10-425-114-30892	Sequence 30892, A
28	66.2	1.4	1485	20	US-10-425-115-111586	Sequence 111586,
29	63.8	1.4	825	17	US-10-369-493-26829	Sequence 26829, A
30	62.6	1.4	1254	17	US-10-438-339-2	Sequence 2, Appli
31	62.6	1.4	1254	18	US-10-416-477-2	Sequence 2, Appli
32	62.6	1.4	1254	21	US-10-754-829A-2	Sequence 2, Appli
33	62.6	1.4	1278	18	US-10-425-114-26294	Sequence 26294, A
34	62.6	1.4	1286	18	US-10-425-114-26222	Sequence 26222, A
35	62.6	1.4	1318	18	US-10-425-114-26279	Sequence 26279, A
c 36	62.6	1.4	2419	17	US-10-108-260A-974	Sequence 974, App
37	62.6	1.4	2626	18	US-10-433-794-32	Sequence 32, Appl
38	62	1.3	4494	18	US-10-263-929-4	Sequence 4, Appli
39	61.6	1.3	3260	9	US-09-858-754-1	Sequence 1, Appli
40	61.6	1.3	3260	19	US-10-728-019-1	Sequence 1, Appli
41	61.6	1.3	5253	9	US-09-858-754-13	Sequence 13, Appl
42	61.6	1.3	5253	13	US-10-000-864-7	Sequence 7, Appli
43	61.6	1.3	5253	19	US-10-728-019-13	Sequence 13, Appl
44	61.4	1.3	1984	18	US-10-425-114-19039	Sequence 19039, A
45	61.4	1.3	2062	18	US-10-425-114-34609	Sequence 34609, A

ALIGNMENTS

RESULT 1

US-09-981-397A-17  
; Sequence 17, Application US/09981397A  
; Publication No. US20030082519A1  
; GENERAL INFORMATION:  
; APPLICANT: Axixma Pharmaceuticals AG  
; APPLICANT: Schubart, Daniel  
; APPLICANT: Habenberger, Peter  
; APPLICANT: Stein-Gerlach, Matthias  
; APPLICANT: Bevec, Dorian  
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their  
; TITLE OF INVENTION: Inhibition  
; FILE REFERENCE: AXM-004.1 US  
; CURRENT APPLICATION NUMBER: US/09/981.397A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR APPLICATION NUMBER: 60/240,750  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-981-397A-17

Query Match 99.6%; Score 4578.4; DB 10; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 5 GGGGGACTGTGCCGTGTGGAACGTGTAGCTGTGGA-AGGTGGACTCTTTTACCATTTGAGG 63

5	Db	 GGGGGACTGTGCGTGTGGAACGTGTAGCTGTGTGAGGCTGACTCTGTATACCATTTGAGG	64
64	Qy	 ATGTTTGGAGGATGAGTATGTGTGGCAGAGGCACACATAAACAGGCAGAGACCCCTTTGCC	123
65	Db	 ATGTTTGGAGATGAGTATGTGTGGCAGAGGCACACATAAACAGGCAGAGACCCCTTTGCC	124
124	Qy	 CTGCGCTTTCTCCCCAACCCCAAGGTGACCTGTGTCTCCAGAGTCTGGAGATTCATAAGT	183
125	Db	 CTGCGCTTTCTCCCCAACCCCAAGGTGACCTGTGTCTCCAGAGTCTGGAGATTCATAAGT	184
184	Qy	 GACCTGCTCTGTGTGTGTGTCTCTCTCAGGATCAGACACAAAGCTGGAGATGGCAGTGATG	243
185	Db	 GACCTGCTCTGTGTGTGTGTCTCTCTCAGGATCAGACACAAAGCTGGAGATGGCAGTGATG	244
244	Qy	 GAAATGGCTTGGCCAGGTGCCCCCTGGCTCAGCAGTGGGCGACGAGAAATCTCCCCAAG	303
245	Db	 GAAATGGCTTGGCCAGGTGCCCCCTGGCTCAGCAGTGGGCGACGAGAAATCTCCCCAAG	304
304	Qy	 CMAAGGAGAACGCGGCCACTGGGGAGAAACAGAGCTCCGTCTTACAAGCTTCAGAGCC	363
305	Db	 CCAAAGGAGAAGCGCGGCCACTGGGGAGAAACAGAGCTCCGTCTTACAAGCTTCAGAGCC	364
364	Qy	 GTGCGAGAAGCGCCTGTGTTCTGCGGAAAAGTGGGAGATCCTGAAATGACGTGATTACCAAG	423
365	Db	 GTGAGAGAGCCTGTGTTCTGCGGAAATGGGAGATCCTGAAATGACGTGATTACCAAG	424
424	Qy	 GGCACAGCCAAAGGAGGCTCCGAGGAGAGGCGCAGCTGCCATCTCTATCATCGCCAGGCT	483
425	Db	 GGCACAGCCAAAGGAGGCTCCGAGGAGAGGCGCAGCTGCCATCTCTATCATCGCCAGGCT	484
484	Qy	 GAGTGTGAGAAATAGCCAAAGATTACGCCACCTTTTCAGAACGCAATTTTCATCGCTGGG	543
485	Db	 GAGTGTGAGAAATAGCCAAAGATTACGCCACCTTTTCAGAACGCAATTTTCATCGCTGGG	544
544	Qy	 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCTCCCAAATGTGGGCCCATGCT	603
545	Db	 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCTCCCAAATGTGGGCCCATGCT	604
604	Qy	 ACAGAGGGCAAAATGGCCGTGTGTGTGGAAGGGAAGCGTCGCAGCAAGAACCCCGAAG	663
605	Db	 ACAGAGGGCAAAATGGCCGTGTGTGTGGAAGGGAAGCGTCGCAGCAAGAACCCCGAAG	664
664	Qy	 AAACGGGAAGAGAGACTCAAAAGTCCCTGGCTCATGCAAGAGTGGCCTTGGCCAAACCC	723
665	Db	 AAACGGGAAGAAGAGACTCAAAAGTCCCTGGCTCATGCAAGAGTGGCCTTGGCCAAACCC	724
724	Qy	 CTCCCAGGACCCCTGAGCAGGAGAGTGCACCATCCCAAGTCAGAGGAGATGAGTCTCCA	783
725	Db	 CTCCCAGGACCCCTGAGCAGGAGAGTGCACCATCCCAAGTCAGAGGAGATGAGTCTCCA	784
784	Qy	 CTCGGCGCCCATATGTTTAGAAAACACCCCGCAGTTTACCAAGGCTCTGAGGAACAGGC	843
785	Db	 CTCGGCGCCCATATGTTTAGAAAACACCCCGCAGTTTACCAAGGCTCTGAGGAACAGGC	844
844	Qy	 CTTGGGCAACTCTGTTTTTAAGAGACTTTGGCGAGGGGCTACGGCGGCTCTGCCCTCGATCA	903
845	Db	 CTTGGGCAACTCTGTTTTTAAGAGACTTTGGCGAGGGGCTACGGCGGCTCTGCCCTCGATCA	904
904	Qy	 GAACTTCACAAACTGATCAGCCCCCTTGCATATGCTGAAACACAGCTGTGGAAATCTGACCAC	963
905	Db	 GAACTTCACAAACTGATCAGCCCCCTTGCATATGCTGAAACACAGCTGTGGAAATCTGACCAC	964
964	Qy	 CCCAGGACGAGGCCCGCTGCCCTGCACCGACCCCTTCCCTTATAGCAGACTGCCT	1023
965	Db	 CCCAGGACGAGGCCCGCTGCCCTGCACCGACCCCTTCCCTTATAGCAGACTGCCT	1024
1024	Qy	 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAAAACCTCACCTCTGGAGTCTTCTCTG	1083
1025	Db	 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAAAACCTCACCTCTGGAGTCTTCTCTG	1084
1084	Qy	 GGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGGCTGACCCACACTTGAGCAAACTG	1143

1085	Db	GGCAAACTGGCCTGTGTAGACAGCCAGAAACCTTGCCTGACCCACACCTGTAGCAAACTG	1144
1144	Qy	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGCCCACTGCCTGTCT	1203
1145	Db	GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGCCCACTGCCTGTCT	1204
1204	Qy	CGTGTGCCCATCAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC	1263
1205	Db	CGTGTGCCCATCAGAAAGTTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC	1264
1264	Qy	GTGAGCTCAAGCCAGGCCCAACGCTGACACGCTTGGCCAAAGACCTTGGGAGACACGGGGC	1323
1265	Db	GTGAGCTCAAGCCAGGCCCAACGCTTGGCCAAAGACCTTGGGAGACACGGGGC	1324
1324	Qy	TCCAGATCCCGGAGCCACGCCCCAAACTGAGCAACCGAGGTGTCTGTCTCACTGAG	1383
1325	Db	TCCAGATCCCGGAGCCACGCCCCAAACTGAGCAACCGAGGTGTCTGTCTCACTGAG	1384
1384	Qy	AAACTCAAGCCAGTGGATTATGATGACCGAGAAAGTCCACTGGGCCACGCAACCACTC	1443
1385	Db	AAACTCAAGCCAGTGGATTATGATGACCGAGAAAGTCCACTGGGCCACGCAACCACTC	1444
1444	Qy	CGCCTGGGAGAGGCTCCTTCGAGAGAGGTGCACAGNATGAGGACACAGCAGACTGGCTTC	1503
1445	Db	CGCCTGGGAGAGGCTCCTTCGAGAGAGGTGCACAGNATGAGGACACAGCAGACTGGCTTC	1504
1504	Qy	CAGTGCCTGTCAAAGAGGTGGCCTGGAAGTATTTTCGGCCAGAGGAGCTGATGGCATGT	1563
1505	Db	CAGTGCCTGTCAAAGAGGTGGCCTGGAAGTATTTTCGGCCAGAGGAGCTGATGGCATGT	1564
1564	Qy	GCAGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGSCCTTGG	1623
1565	Db	GCAGATTGACCTCACCCAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGSCCTTGG	1624
1624	Qy	GTCAACATCTTCATGAGACTGCTGGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGACAG	1683
1625	Db	GTCAACATCTTCATGAGACTGCTGGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGACAG	1684
1684	Qy	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTTGGAGGGTCTGGAA	1743
1685	Db	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTTGGAGGGTCTGGAA	1744
1744	Qy	TACCTTCACTCACAAAGATTCATGAGGGAGCTCAAAAGCTGACAAGCTGCTCCTCTGTCC	1803
1745	Db	TACCTTCACTCACAAAGATTCATGAGGGAGCTCAAAAGCTGACAAGCTGCTCCTCTGTCC	1804
1804	Qy	AGCGATGGAGCCACGACGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT	1863
1805	Db	AGCGATGGAGCCACGACGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT	1864
1864	Qy	GGCTTGGGAAAAGTCTTGTCTCACAGGGGACTACATCCCTGGCAAGAGCCCAATGCT	1923
1865	Db	GGCTTGGGAAAAGTCTTGTCTCACAGGGGACTACATCCCTGGCAAGAGCCCAATGCT	1924
1924	Qy	CCGAGAGTGGTCTGGGCAGAGCTGCGACGCCAAGGTGGATGTCTGAGCAGCTGCTGT	1983
1925	Db	CCGAGAGTGGTCTGGGCAGAGCTGCGACGCCAAGGTGGATGTCTGAGCAGCTGCTGT	1984
1984	Qy	ATGATGCTGCACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC	2043
1985	Db	ATGATGCTGCACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC	2044
2044	Qy	TGCCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGCGCCCT	2103
2045	Db	TGCCTCAAGATTGCCAGCGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGCGCCCT	2104
2104	Qy	CTCACAGCCCAAGGCCATCCAAAGAGGGGTGAGGAAAGGCCAATCCAACGGGTGTCTGCA	2163
2105	Db	CTCACAGCCCAAGGCCATCCAAAGAGGGGTGAGGAAAGGCCAATCCAACGGGTGTCTGCA	2164
2164	Qy	GGGAGTGGGAGGAAAGGTGAAACGGGCACTACAGCAAGTGGGAGGTTCTGAAGAGCCCT	2223
2165	Db	GGGAGTGGGAGGAAAGGTGAAACGGGCACTACAGCAAGTGGGAGGTTCTGAAGAGCCCT	2224



QY 2224 TGGAGGGGAGAAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAATTATACCACGAG 2283  
DB |||||  
2225 TGGAGGGGAGAAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAATTATACCACGAG 2284  
QY |||||  
2284 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGGCCACAGGGCCCGCGGACAGCTGAG 2343  
DB |||||  
2285 ACCCTCCATGCCAGCCGAGAGAGCTTTGCGCAAGGGGCCACAGGGCCCGCGGACAGCTGAG 2344  
QY |||||  
2344 GAGACAAACAGCAGAGAGCCCTTAAGCTCCAGGCTCTCTCTCCACAGAGCCGCCACAGAGCCA 2403  
DB |||||  
2345 GAGACAAACAGCAGAGAGCCCTTAAGCTCCAGGCTCTCTCTCCACAGAGCCGCCACAGAGCCA 2404  
QY |||||  
2404 AACAAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAAACCCCTTACCT 2463  
DB |||||  
2405 AACAAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAAACCCCTTACCT 2464  
QY |||||  
2464 CTGTCTCTCTGGAGCCAGCCCTTGGCAGAAACCCAGACTCACAGAGCGGAAAGCAACC 2523  
DB |||||  
2465 CTGTCTCTCTGGAGCCAGCCCTTGGCAGAAACCCAGACTCACAGAGCGGAAAGCAACC 2524  
QY |||||  
2524 GTCCCGGAGCAGGAATCGCAGAGTGGAAATAGAAATTTATTTCTCAACAGAGCTGTCCAG 2583  
DB |||||  
2525 GTCCCGGAGCAGGAATCGCAGAGTGGAAATAGAAATTTATTTCTCAACAGAGCTGTCCAG 2584  
QY |||||  
2584 CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCGACAGGCTCTCC 2643  
DB |||||  
2644 CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCGACAGGCTCTCC 2644  
QY |||||  
2644 CTGTCTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGTCTCGGGGACACCCCTG 2703  
DB |||||  
2645 CTGTCTGGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGGTCTCGGGGACACCCCTG 2704  
QY |||||  
2704 AGCTCAGGCTACACTCTCTGAGCAGCCAGGCTCGAAGCTCGAAGCTCGGAGTGAACATG 2763  
DB |||||  
2705 AGCTCAGGCTACACTCTCTGAGCAGCCAGGCTCGAAGCTCGAAGCTCGGAGTGAACATG 2764  
QY |||||  
2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGTAAGTCCAA 2823  
DB |||||  
2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGTAAGTCCAA 2824  
QY |||||  
2824 ATACAGTCTCTTAATGTGTGAACACTGTCACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2883  
DB |||||  
2883 GACATCGCCACTGGGCATCAGCAGCCAGATCCAGCTCGAGCTTCAGCTTGGTCAACCAA 2943  
DB |||||  
2943 GACATCGCCACTGGGCATCAGCAGCCAGATCCAGCTCGAGCTTCAGCTTGGTCAACCAA 2944  
QY |||||  
2944 GACGGGCAAGCTGTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGC 3003  
DB |||||  
2945 GACGGGCAAGCTGTTCGCTACGACATGGAGGTGCCAGACTCGGGCATCGACCTGCAGTGC 3004  
QY |||||  
3004 ACATGGCCCTGTATGGCAGCTTCCGCTGGAGCTGGAGGGTCAAGCATATGGCCAGCTGGAG 3063  
DB |||||  
3063 ACATGGCCCTGTATGGCAGCTTCCGCTGGAGCTGGAGGGTCAAGCATATGGCCAGCTGGAG 3064  
QY |||||  
3064 AACAGGCCCTAACCTTGCCCTCCACCGCGGCTCCACACTCGCGGAAAGAGCCCTTCTGT 3123  
DB |||||  
3123 AACAGGCCCTAACCTTGCCCTCCACCGCGGCTCCACACTCGCGGAAAGAGCCCTTCTGT 3124  
QY |||||  
3124 CTCGGTGACGATGTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGGCC 3183  
DB |||||  
3125 CTCGGTGACGATGTGCTGCTGAAACACAGGCTCAGCGCTTCCAGGGGATTTGCCAGGCC 3184  
QY |||||  
3184 CCCCGGCTTACAGTGGGAAACAGGGCTTCGAGCAGCAAGTGGGGGCAACAGAAATGCC 3243  
DB |||||  
3185 CCCCGGCTTACAGTGGGAAACAGGGCTTCGAGCAGCAAGTGGGGGCAACAGAAATGCC 3244  
QY |||||  
3244 TCCAGGATTTTACACCTGAGCCCTGCCCCACCTGTGTAAGAAACATCCGCCACGTGAA 3303  
DB |||||  
3245 TCCAGGATTTTACACCTGAGCCCTGCCCCACCTGTGTAAGAAACATCCGCCACGTGAA 3304

QY 3304 GAGACAGAGAGGATGGCAGGAGTTACTCTGGGGAACAAACAGGGATCTTTTCTTGCC 3363  
DB |||||  
3305 GAGACAGAGAGGATGGCAGGAGTTACTCTGGGGAACAAACAGGGATCTTTTCTTGCC 3364  
QY |||||  
3364 CCTGCTCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423  
DB |||||  
3365 CCTGCTCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424  
QY |||||  
3424 GGAGAGCAGCTTCCAGGCTGGGTGAGAGGGGTGGGAGAGCCCTTCGGGCCCTCACCCCTC 3483  
DB |||||  
3425 GGAGAGCAGCTTCCAGGCTGGGTGAGAGGGGTGGGAGAGCCCTTCGGGCCCTCACCCCTC 3484  
QY |||||  
3484 CAGGCTGCTGAGAGTGTCAAGTGTAAAGGGCCCAAACTCAGGTTTCAGTGACGAGAACCA 3543  
DB |||||  
3485 CAGGCTGCTGAGAGTGTCAAGTGTAAAGGGCCCAAACTCAGGTTTCAGTGACGAGAACCA 3544  
QY |||||  
3544 GGTTCAGCAGGTATGCCCGCGGTAGGTTAAAGGGGGCCCTCTAAACCCCTTCGCTGGCCCTC 3603  
DB |||||  
3545 GGTTCAGCAGGTATGCCCGCGGTAGGTTAAAGGGGGCCCTCTAAACCCCTTCGCTGGCCCTC 3604  
QY |||||  
3604 ACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAAGATGCCTGACCCCTGGGAAGGCT 3663  
DB |||||  
3605 ACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAAAGATGCCTGACCCCTGGGAAGGCT 3664  
QY |||||  
3664 CCCTGGTAGAATACACACACTTTTCAAGTTTGTGCAACACAGGTCCTGAGTTGACCTCT 3723  
DB |||||  
3723 CCCTGGTAGAATACACACACTTTTCAAGTTTGTGCAACACAGGTCCTGAGTTGACCTCT 3724  
QY |||||  
3724 GGTTCAGCAGGACCAAGAGGTGTAAAGTGAAGTGTCTCAGTCCGCCAGACATGT 3783  
DB |||||  
3725 GGTTCAGCAGGACCAAGAGGTGTAAAGTGAAGTGTCTCAGTCCGCCAGACATGT 3784  
QY |||||  
3784 GCGCTTTGTCTGTGGCTTACCACTTTCCCAGACAGCAGGCGCCCGAGCCCTTCAGGC 3843  
DB |||||  
3785 GCGCTTTGTCTGTGGCTTACCACTTTCCCAGACAGCAGGCGCCCGAGCCCTTCAGGC 3844  
QY |||||  
3844 CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGGTGTAT 3903  
DB |||||  
3845 CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGGTGTAT 3904  
QY |||||  
3904 GCAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAGGAAAGCAG 3963  
DB |||||  
3905 GCAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAGGAAAGCAG 3964  
QY |||||  
3964 CAGAGAGAGCTCCGCGGCCCGAGCCCACTATCAGTGTCCAGCTGCTGGTTCGCCAG 4023  
DB |||||  
3965 CAGAGAGAGCTCCGCGGCCCGAGCCCACTATCAGTGTCCAGCTGCTGGTTCGCCAG 4024  
QY |||||  
4024 AGCACAGCTCAGCATCACACTGACACTCACCTTCCCTGCGCCCTGGCCAGAGGGTACTGC 4083  
DB |||||  
4083 AGCACAGCTCAGCATCACACTGACACTCACCTTCCCTGCGCCCTGGCCAGAGGGTACTGC 4084  
QY |||||  
4084 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCGCTCTGGCAGGGC 4143  
DB |||||  
4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCGCTCTGGCAGGGC 4144  
QY |||||  
4144 AGGSCAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGAGACA 4203  
DB |||||  
4145 AGGSCAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGAGACA 4204  
QY |||||  
4204 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGAAATA 4263  
DB |||||  
4205 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGAAATA 4264  
QY |||||  
4264 GCACCCCTTCCCATTTGCCAGCTTACGCTCTTCCCGAGGGAGGACTACTCAGGACTC 4323  
DB |||||  
4265 GCACCCCTTCCCATTTGCCAGCTTACGCTCTTCCCGAGGGAGGACTACTCAGGACTC 4324  
QY |||||  
4324 ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGACCTCAACCTCTGGG 4383  
DB |||||  
4325 ACGTAGCATTTAAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCTCAACCTCTGGG 4384  
QY |||||  
4384 GCAGGGGACGCCGAGACTCGTGGGAGAAAGCTCATTTCCCACTCTTGGCAAGACAGGCTT 4443

Db 4385 GCAGGGACCGCAGACTCCGTGGGAGAAGCTCAITCCACATCTTGCCAAACAGACGCTT 4444  
Qy 4444 TGTCAGCTGTCCACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAC 4503  
Db 4445 TGTCAGCTGTCCACATTGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAC 4504  
Qy 4504 ATAAAGAACTGCAGCCTTGTTGTTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4563  
Db 4505 ATAAAGAACTGCAGCCTTGTTGTTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4564  
Qy 4564 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4595  
Db 4565 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4596

RESULT 2

US-10-172-118-975  
; Sequence 975, Application US/10172118  
; Publication No. US20030224374A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veer, Laura  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172,118  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 975  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: NM\_003954  
; DATABASE ENTRY DATE: 2001-06-18  
US-10-172-118-975

Query Match 99.6%; Score 4578.4; DB 17; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 5 GGGGACTGTGCGGTGGAACTGTAGCTGTGA-AGGTGACTCTGTACATTGAGG 63  
Db 5 GGGGACTGTGCGGTGGAACTGTAGCTGTGTAGAGGTGACTCTGTACATTGAGG 64  
Qy 64 ATGTTTGGAGTAGTATGTGTGGCAGAGGCACATATAACAGCAGAGACCTTTGCC 123  
Db 65 ATGTTTGGAGTAGTATGTGTGGCAGAGGCACATATAACAGCAGAGACCTTTGCC 124  
Qy 124 CTGCTCTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGCTCTGGGATTTCAAGT 183  
Db 125 CTGCTCTTCTCCCAACCCAGGCTGACCTGTGTTCTCCAGCTCTGGGATTTCAAGT 184  
Qy 184 GACCTGCTCTGTGTTGTTCTCTCTCAGGATGAGCACAAGCTCTGGAGATGGCAGTGATG 243  
Db 185 GACCTGCTCTGTGTTGTTCTCTCTCAGGATGAGCACAAGCTCTGGAGATGGCAGTGATG 244  
Qy 244 GAAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGACAGAGAACTCCCCAAG 303  
Db 245 GAAATGGCTGCCAGGTGCCCTCGCTCAGCAGTGGGACAGAGAACTCCCCAAG 304  
Qy 304 CCAAGGAGAGACGCCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGAGGCC 363  
Db 305 CCAAGGAGAGACGCCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGAGGCC 364

Qy 364 GTGAGAAAGAGCCCTGTGTCTTGGGAAAGTGGAGATCTCTCAATGACGTGATTACCAAG 423  
Db GTGAGAAAGAGCCCTGTGTCTTGGGAAAGTGGAGATCTCTCAATGACGTGATTACCAAG 424  
Qy 424 GGCAAGCAAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483  
Db GGCAAGCAAGGAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484  
Qy 484 GAGTGTGAGATAGCCAAAGATTGAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 543  
Db GAGTGTGAGATAGCCAAAGATTGAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 544  
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCATGCT 603  
Db TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCATGCT 604  
Qy 604 ACAGAGGCAAAATGGCCCGTGTGTTGGAAAGGAAAGCGTTCGAGCAAAAGCCCGGAAG 663  
Db ACAGAGGCAAAATGGCCCGTGTGTTGGAAAGGAAAGCGTTCGAGCAAAAGCCCGGAAG 664  
Qy 664 AAACGGAAGAAAGAGCTCAAAAGTCCCTGCTCATGCAAGAGTGGCCCTTGSCCAAAACC 723  
Db AAACGGAAGAAAGAGCTCAAAAGTCCCTGCTCATGCAAGAGTGGCCCTTGSCCAAAACC 724  
Qy 724 CTCCCCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGATGAGTCTCCA 783  
Db CTCCCCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGATGAGTCTCCA 784  
Qy 784 CTCGGCGCCCATATGTTTAGAAACACCCCGCAGTTTACCAAGCTCTGAAAGAAACAGGC 843  
Db CTCGGCGCCCATATGTTTAGAAACACCCCGCAGTTTACCAAGCTCTGAAAGAAACAGGC 844  
Qy 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGAGGCTTACGGCGGCTTGCCTCGATCA 903  
Db CTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGAGGCTTACGGCGGCTTGCCTCGATCA 904  
Qy 904 GAACTCCACAAACTGATCAGGCCCTTGCAATGTCTGAAACCACTGTGGAACATGCACCAC 963  
Db GAACTCCACAAACTGATCAGGCCCTTGCAATGTCTGAAACCACTGTGGAACATGCACCAC 964  
Qy 964 CCCAGGACGAGGCGCCCTGCCCCGTGCCACAGCACCCCTTCCCTATAGCAGACTGCT 1023  
Db CCCAGGACGAGGCGCCCTGCCCCGTGCCACAGCACCCCTTCCCTATAGCAGACTGCT 1024  
Qy 1024 CATCCCTTCCCATCCACCTCTCAGCCCTGGAACCTCACCCTCTCGAGTCTTCTCTG 1083  
Db CATCCCTTCCCATCCACCTCTCAGCCCTGGAACCTCACCCTCTCGAGTCTTCTCTG 1084  
Qy 1084 GGCACAACTGGCTGTGTAGACAGCAGCAAAACCCCTTGGCTGACCCACACCTGAGCAAACTG 1143  
Db GGCACAACTGGCTGTGTAGACAGCAGCAAAACCCCTTGGCTGACCCACACCTGAGCAAACTG 1144  
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGAGGCCAGCTGCTGTCT 1203  
Db GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGAGGCCAGCTGCTGTCT 1204  
Qy 1204 CGTGTGCCCAGTGAAGTTTTCTGTGAGAGAAATACCTAGTGCATGCTCTGCAAGCGAGC 1263  
Db CGTGTGCCCAGTGAAGTTTTCTGTGAGAGAAATACCTAGTGCATGCTCTGCAAGCGAGC 1264  
Qy 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACCGCTGGCCAAAGACCTGGGACAGCAGGGGC 1323  
Db GTGAGCTCAAGCCAGGCCCCACAGCTGACCGCTGGCCAAAGACCTGGGACAGCAGGGGC 1324  
Qy 1324 TCCAGATCCCGGAGCCCCAGCCCCAAAACCTGAGGACAAAGAGGTGTCTGCTCACTGAG 1383  
Db TCCAGATCCCGGAGCCCCAGCCCCAAAACCTGAGGACAAAGAGGTGTCTGCTCACTGAG 1384  
Qy 1384 AAACCTCAAGCCAGTGGATTATGAGTACCGAGAGAAAGTCCACTGGGCGCACGCCAGCTC 1443  
Db AAACCTCAAGCCAGTGGATTATGAGTACCGAGAGAAAGTCCACTGGGCGCACGCCAGCTC 1444  
Qy 1444 GCCTGGGACAGAGGCTCTCTCGGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1503

Db	1445		CGCTGGGACAGAGCTCTCTCGAGAGGTGCACAGGATGGAGGACAAGCAGACTGGCTTC	1504
Qy	1504		CAGTGGCGCTGTCAAAAAGGTGGCGCTGGAAGTATTTTCGGGACAGAGGAGCTGATGGCATGT	1563
Db	1505		CAGTGGCGCTGTCAAAAAGGTGGCGCTGGAAGTATTTTCGGGACAGAGGAGCTGATGGCATGT	1564
Qy	1564		GCAGGATTGACCTCACCCAGAAATTTGTCCTTTGATGGAGCTGTGAGAGAGGGCCCTTGG	1623
Db	1565		GCAGGATTGACCTCACCCAGAAATTTGTCCTTTGATGGAGCTGTGAGAGAGGGCCCTTGG	1624
Qy	1624		GTCAACATCTCATGGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGGAGCAG	1683
Db	1625		GTCAACATCTCTCATGGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGCTGTGTCAAGGAGCAG	1684
Qy	1684		GGCTGTCTCCACAGAGGACCGGGCCCTGTACTACCTGGGCCAGAGCCCTGGAGGGTCTCGAA	1743
Db	1685		GGCTGTCTCCACAGAGGACCGGGCCCTGTACTACCTGGGCCAGAGCCCTGGAGGGTCTCGAA	1744
Qy	1744		TACTCCACTCAAGAGGATTTCTGCATGGGACGTCAAAGCTGACAAAGTGTCTCTCTGTC	1803
Db	1745		TACTCCACTCAAGAGGATTTCTGCATGGGACGTCAAAGCTGACAAAGTGTCTCTCTGTC	1804
Qy	1804		AGCGATGGAGGCCACGAGCCCTCTGTGAATTTGGGCCATGTGTGTCTTCAACCTGAT	1863
Db	1805		AGCGATGGAGGCCACGAGCCCTCTGTGAATTTGGGCCATGTGTGTCTTCAACCTGAT	1864
Qy	1864		GGCTTGGGAAAGTCTTGTCTCACAGGGGACTACATCCCTGSCACAGAGACCCACATGGCT	1923
Db	1865		GGCTTGGGAAAGTCTTGTCTCACAGGGGACTACATCCCTGSCACAGAGACCCACATGGCT	1924
Qy	1924		CCGAGGTGTGTCTGGGACGAGCTGGGACCCCAAGGTGATGTCTGGAGCAGCTGTGT	1983
Db	1925		CCGAGGTGTGTCTGGGACGAGCTGGGACCCCAAGGTGATGTCTGGAGCAGCTGTGT	1984
Qy	1984		ATGATGTGCACATGTCTCAAGCGTGTGCCACCCCTGGAATCTAGTTCTTCGAGGGCCGCTC	2043
Db	1985		ATGATGTGCACATGTCTCAAGCGTGTGCCACCCCTGGAATCTAGTTCTTCGAGGGCCGCTC	2044
Qy	2044		TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGCCCT	2103
Db	2045		TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGCCCT	2104
Qy	2104		CTCAGCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA	2163
Db	2105		CTCAGCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA	2164
Qy	2164		GCGAGCTGGAGGGAAAGTGAAACCGGCACTACAGCAAGTGGAGGTCTGAAAGAGCCCT	2223
Db	2165		GCGAGCTGGAGGGAAAGTGAAACCGGCACTACAGCAAGTGGAGGTCTGAAAGAGCCCT	2224
Qy	2224		TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG	2283
Db	2225		TGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG	2284
Qy	2284		ACCTCTCATGCCAGCGAGAGAGCTTTCGCAAGGGCCCGAGGGCCCGGCGCAGCTGAG	2343
Db	2285		ACCTCTCATGCCAGCGAGAGAGCTTTCGCAAGGGCCCGAGGGCCCGGCGCAGCTGAG	2344
Qy	2344		GAGACACAGGACAGAGCCCTTAAGCTCCAGCTCTCTTCCACAGAGCCCGCCACAGGCCA	2403
Db	2345		GAGACACAGGACAGAGCCCTTAAGCTCCAGCTCTCTTCCACAGAGCCCGCCACAGGCCA	2404
Qy	2404		AACAAGTCTCTCTCTTGTACTTTGAGCAAGAGAGGTCTGGGATGTGGGAAACCTTACCT	2463
Db	2405		AACAAGTCTCTCTCTTGTACTTTGAGCAAGAGAGGTCTGGGATGTGGGAAACCTTACCT	2464
Qy	2464		CTGTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2523
Db	2465		CTGTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAGCAACC	2524
Qy	2524		GTCCCGGAGCAGGAATGCAGCAGCTGGAAATAGAAATTTATTCCTCAACAGCCTGTCCAG	2583

Db	2525		GTCCCGGAGCAGGAATGTCAGCAGCTGGAAATAGAAATTTATTCCTCAACAGCCTGTCCAG	2584
Qy	2584		CCATTTTCTCTGGAGGACAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCC	2643
Db	2585		CCATTTTCTCTGGAGGACAGGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCTCTCC	2644
Qy	2644		CTGTCCGATCAGCTGAGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCTG	2703
Db	2645		CTGTCCGATCAGCTGAGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCTG	2704
Qy	2704		AGCTCAGCGCTACACTCTCGAGCAGCAGCGGAGGCTCGAAGCTCAGCTCGGAAATG	2763
Db	2705		AGCTCAGCGCTACACTCTCGAGCAGCAGCGGAGGCTCGAAGCTCAGCTCGGAAATG	2764
Qy	2764		GTCTGSCCGCGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGAAAGTCCAA	2823
Db	2765		GTCTGSCCGCGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGAAAGTCCAA	2824
Qy	2824		ATACAGTCTCTTAATGTGTAACACTCGGCATCGGCAGTTCACCGGGTCAAAAGTGGGA	2883
Db	2825		ATACAGTCTCTTAATGTGTAACACTCGGCATCGGCAGTTCACCGGGTCAAAAGTGGGA	2884
Qy	2884		GACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTCAGCTTGGTCAACAA	2943
Db	2885		GACATCGCCACTGGCATCAGCAGCCAGATCCCAAGCTTCAGCTTGGTCAACAA	2944
Qy	2944		GACGGCAGCTGTTCCTACGACATGAGGTGTCAGACTCGGGCATCGACCTCGAGTGC	3003
Db	2945		GACGGCAGCTGTTCCTACGACATGAGGTGTCAGACTCGGGCATCGACCTCGAGTGC	3004
Qy	3004		ACACTGSCCCTGTATGTCAGCTTGCCTGAGCTGAGGTCAAGCATGAGCAGCTGGAG	3063
Db	3005		ACACTGSCCCTGTATGTCAGCTTGCCTGAGCTGAGGTCAAGCATGAGCAGCTGGAG	3064
Qy	3064		AACAGGCCCTTAAACCTGSCCTCCACCGCCGCTCCACACTGCGGAAAGCAGCTTCTG	3123
Db	3065		AACAGGCCCTTAAACCTGSCCTCCACCGCCGCTCCACACTGCGGAAAGCAGCTTCTG	3124
Qy	3124		CTCGTGCAGATGTCTCCCTGAAAAACAGAGGTTCAGCTTCCAGGGATTCAGCAGCC	3183
Db	3125		CTCGTGCAGATGTCTCCCTGAAAAACAGAGGTTCAGCTTCCAGGGATTCAGCAGCC	3184
Qy	3184		CCCCGGCTCAGTGGGAAACAGGGCTTCGACGACGCAAGTGGGGGCAAGCAATGACC	3243
Db	3185		CCCCGGCTCAGTGGGAAACAGGGCTTCGACGACGCAAGTGGGGGCAAGCAATGACC	3244
Qy	3244		TCCAGGATTTACACTGAGCCCTGCCCCACCTGCTGAAAAAACAATCCGCCACGTGAA	3303
Db	3245		TCCAGGATTTACACTGAGCCCTGCCCCACCTGCTGAAAAAACAATCCGCCACGTGAA	3304
Qy	3304		GAGACAGGAGGATGGCAGGAGTTACCTGGGAAAAACAAACAGGGATCTTTTCTGCCC	3363
Db	3305		GAGACAGGAGGATGGCAGGAGTTACCTGGGAAAAACAAACAGGGATCTTTTCTGCCC	3364
Qy	3364		CCTGCTCCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG	3423
Db	3365		CCTGCTCCAGTTCGAGTTGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG	3424
Qy	3424		GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGGAGCCCTTCGGCCCTCACCCCTC	3483
Db	3425		GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGGAGCCCTTCGGCCCTCACCCCTC	3484
Qy	3484		CAGGCTGCTGTGAGAGTCAAGTGTGTAAAGGGCCCAAACTCAGGTTTCAGTGCAGAACCA	3543
Db	3485		CAGGCTGCTGTGAGAGTCAAGTGTGTAAAGGGCCCAAACTCAGGTTTCAGTGCAGAACCA	3544
Qy	3544		GGTCAGCAGGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGGCCTC	3603
Db	3545		GGTCAGCAGGTATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGGCCTC	3604
Qy	3604		ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAAAGATGCTGACCTTGGGAAAGGCT	3663
Db	3605		ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAAAGATGCTGACCTTGGGAAAGGCT	3664

Qy	3664	CCCTGGTAGAATA	CACACACATTTTTCAGGTTGTTGGCAACACAGAGTCTCTGAGTTGACCTCT	3723
Db	3665	CCCTGGTAGAATA	CACACACATTTTTCAGGTTGTTGGCAACACAGGTCCTGAGTTGACCTCT	3724
Qy	3724	GGTTCCAGCAGGACCA	AAAGAGGTGTGAAGTGAAGTGGTTCTCAGTCCCCAGACATGT	3783
Db	3725	GGTTCCAGCAGGACCA	AAAGAGGTGTGAAGTGAAGTGGTTCTCAGTCCCCAGACATGT	3784
Qy	3784	GCCCTTTGCTGCTGG	GTATCCCTTTCCACAGACAGCAGGCCCCCGAGCCCTTTCAGGC	3843
Db	3785	GCCCTTTGCTGCTGG	GTATCCCTTTCCACAGACAGCAGGCCCCCGAGCCCTTTCAGGC	3844
Qy	3844	CCAGCACTGCCCCAG	ACTCGCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGGTGAT	3903
Db	3845	CCAGCACTGCCCCAG	ACTCGCTGGCACTCAGTTCCTCATCTGTGTAAAGGTGAAGGGTGAT	3904
Qy	3904	GCAGGATATGCTCAT	CACAGGAACACTCTGTGTGATGGACATGATCAGTCTTAAGGAAGCAG	3963
Db	3905	GCAGGATATGCTCAT	CACAGGAACACTCTGTGTGATGGACATGATCAGTCTTAAGGAAGCAG	3964
Qy	3964	CAGAGAGAGAGTCT	CGGGCGCCCCAGCCCCCATATCAGTGTCCAGCGTCTGTTTCCCAG	4023
Db	3965	CAGAGAGAGAGTCT	CGGGCGCCCCAGCCCCCATATCAGTGTCCAGCGTCTGTTTCCCAG	4024
Qy	4024	AGCAAGCTCAGCAT	ATCACTGACATCAACCTGCCCTGCCCTGGCCAGAGGGTACTGC	4083
Db	4025	AGCAAGCTCAGCAT	ATCACTGACATCAACCTGCCCTGCCCTGGCCAGAGGGTACTGC	4084
Qy	4084	CGAGGCACTTTGC	ACTCTGATGACCTTCAAGCACTTTTCATGGCTGCCCTCTGCGAGGC	4143
Db	4085	CGAGGCACTTTGC	ACTCTGATGACCTTCAAGCACTTTTCATGGCTGCCCTCTGCGAGGC	4144
Qy	4144	AGGGCAGGGCAGT	GACACTGTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGAACA	4203
Db	4145	AGGGCAGGGCAGT	GACACTGTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGAACA	4204
Qy	4204	GTCTTGAGCTGTCC	ACATGCAATGTGACTCTCAAACCTCTTTCACAGATTTCTCTAAGAATA	4263
Db	4205	GTCTTGAGCTGTCC	ACATGCAATGTGACTCTCAAACCTCTTTCACAGATTTCTCTAAGAATA	4264
Qy	4264	GCACCCCTTCCCAT	TTCGCCACGTTAGCCTCTTCCAGGGGAGCTACTCAGGACTC	4323
Db	4265	GCAACCCCTTCCCAT	TTCGCCACGTTAGCCTCTTCCAGGGGAGCTACTCAGGACTC	4324
Qy	4324	ACGTAGCATTTAAAT	CAGCTGTGAATCGTCAGGGGTGTCTGTAGCCTCAACCTCTCTGGG	4383
Db	4325	ACGTAGCATTTAAAT	CAGCTGTGAATCGTCAGGGGTGTCTGTGTAGCCTCAACCTCTCTGGG	4384
Qy	4384	GCAGGGAGCGCCGA	AGACTCTCCCTGGGAGAGTCTATTCACATCTTGTGCAAGACAGCCTT	4443
Db	4385	GCAGGGAGCGCCGA	AGACTCTCCCTGGGAGAGTCTATTCACATCTTGTGCAAGACAGCCTT	4444
Qy	4444	TGTCAGCTGTCCAT	TGAGTGCAGCTGCTCCGGGGAGAGCCCGCGCCCCCAGCAC	4503
Db	4445	TGTCCAGCTGTCCAT	TGAGTGCAGCTGCTCCGGGGAGAGCCCGCGCCCCCAGCAC	4504
Qy	4504	ATAAAGAACTCG	ACCTTGGTACTGTCAGAGTCTGGGTGTGTAGAACTCTTTGTAAAGCA	4563
Db	4505	ATAAAGAACTCG	ACGCTTGGTACTGTCAGAGTCTGGGTGTGTAGAACTCTTTGTAAAGCA	4564
Qy	4564	TAAAGTTTGGGGT	GATGACAAATGTTTAAAAA	4595
Db	4565	TAAAGTTTGGGGT	GATGACAAATGTTTAAAAA	4596

RESULT 3  
US-10-342-887-975  
; Sequence 975, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong

```

? APPLICANT: Linsley, Peter S.
? APPLICANT: Mao, Mao
? APPLICANT: Roberts, Christopher J.
? APPLICANT: Van 't Veer, Laura Johanna
? APPLICANT: Van de Vijver, Marc J.
? APPLICANT: Bernards, Rene
? TITLE OF INVENTION: diagnosis and Prognosis of Breast Cancer Patients
? FILE REFERENCE: 9301-188-999
? CURRENT APPLICATION NUMBER: US/10/342,887
? CURRENT FILING DATE: 2003-01-15
? PRIOR APPLICATION NUMBER: 60/298,918
? PRIOR FILING DATE: 2001-06-18
? PRIOR APPLICATION NUMBER: 60/380,710
? PRIOR FILING DATE: 2002-05-14
? PRIOR APPLICATION NUMBER: 10/172,118
? PRIOR FILING DATE: 2002-06-14
? NUMBER OF SEQ ID NOS: 2699
? SEQ ID NO 975
? LENGTH: 4596
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-342-887-975

```

Query Match	99.6%;	Score 4578.4;	DB 18;	Length 4596;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4590;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
Qy	5	GGGGGACTGTGCGGTGTGGAAACGTGTAGCTGTTCGA-AGGTGGACTCTCTTACCAATTGAGG	63	
Db	5	GGGGGACTGTGCGGTGTGGAAACGTGTAGCTGTTCGAGAGGTGGACTCTGTATTACCAATTGAGG	64	
Qy	64	ATGTTTTCGAGGATGAGTATGTGTGGCAGAGGSCACACATAAACGGCAGAGACCCCTTTTGCC	123	
Db	65	ATGTTTTCGAGGATGAGTATGTGTGGCAGAGGSCACATAAACGGCAGAGACCCCTTTTGCC	124	
Qy	124	CCTGCTTTTCTCCCCAACCAACGAGCTGACCTGTGTTCCTCCAGGTCCTGGGATTTCTAAGT	181	
Db	125	CCTGCTTTTCTCCCCAACCAACGAGCTGACCTGTGTTCCTCCAGGTCCTGGGATTTCTAAGT	184	
Qy	184	GACCTGCTGTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCCTGGGAGATGGCAGTGTATG	243	
Db	185	GACCTGCTGTGTGTTTGGTCTCTCTCAGGATGAGCACAAAGCCTGGGAGATGGCAGTGTATG	244	
Qy	244	GAATGCGCTGCCAGGTGCCCTGGCTCAGCACTGGGCGAGCAGAGGAAGTCTCCCCAAG	303	
Db	245	GAATGCGCTGCCAGGTGCCCTGGCTCAGCACTGGGCGAGCAGAGGAAGTCTCCCCAAG	304	
Qy	304	CCAAAGGAGAAAGACCGCCCACTGGGGAAGAAACAGAGCTCCGTCACAAAGCTTGAGGCC	363	
Db	305	CCAAAGGAGAAAGACCGCCCACTGGGGAAGAAACAGAGCTCCGTCACAAAGCTTGAGGCC	364	
Qy	364	GTGAGAAAGAGCCCTGTGTCTGCGGAAAGTGGGAGATCTCTGAATGACGTGATTACCAAG	423	
Db	365	GTGAGAAAGAGCCCTGTGTCTGCGGAAAGTGGGAGATCTCTGAATGACGTGATTACCAAG	424	
Qy	424	GGCACAGCCAAAGAAAGGCTCCGAGGCAAGGCGACGTGCAATCTCTATCATCGCCAGGCT	483	
Db	425	GGCACAGCCAAAGAAAGGCTCCGAGGCAAGGCGACGTGCAATCTCTATCATCGCCAGGCT	484	
Qy	484	GAGTGTGAGAAATAGCCAAAGATTGAGCCCACTTTTCAGAACGCAATTTTTCATCGCTGGG	543	
Db	485	GAGTGTGAGAAATAGCCAAAGATTGAGCCCACTTTTCAGAACGCAATTTTTCATCGCTGGG	544	
Qy	544	TCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGCCCCATGCT	603	
Db	545	TCCAAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGCCCCATGCT	604	
Qy	604	ACAGAGGCGAAAATGGGCCGTGTGTGTTGGAAGGGAAGCGTCGCAGCAAAAGCCCGGAAG	663	
Db	605	ACAGAGGCGAAAATGGGCCGTGTGTGTTGGAAGGGAAGCGTCGCAGCAAAAGCCCGGAAG	664	
Qy	664	AAACGGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGCAAGAGTGGCTTTGGCCAAACCC	723	

Db 665 AAACGGAAAGAGAGAGCTCAAAGTCCCTGGCTCATGACAGGAGTGCGCTTGGCCAAACCC 724  
Qy 724 CTCCCAGAGACCCCTGAGCAGGAGAGCTGCACCACTCCAGTGCAGAGGAGTGAAGTCTCCA 783  
Db 725 CTCCCAGAGACCCCTGAGCAGGAGAGCTGCACCACTCCAGTGCAGAGGAGTGAAGTCTCCA 784  
Qy 784 CTGGCGGCCCATATGTTAGAAAACACCCCGCAGTTTACCAAGACCTCTGAAAGAAACCAAGGC 843  
Db 785 CTGGCGGCCCATATGTTAGAAAACACCCCGCAGTTTACCAAGACCTCTGAAAGAAACCAAGGC 844  
Qy 844 CTGGCGAACCTCTGTTTAAAGCAGCTTGGCAGGAGGCTTACCGCGGCTCTGCCTCGATCA 903  
Db 845 CTGGCGAACCTCTGTTTAAAGCAGCTTGGCAGGAGGCTTACCGCGGCTCTGCCTCGATCA 904  
Qy 904 GAACTCCACAACTGATCAGCCCTTGCATATGTTGAACACAGTGTGAAACTGCACAC 963  
Db 905 GAACTCCACAACTGATCAGCCCTTGCATATGTTGAACACAGTGTGAAACTGCACAC 964  
Qy 964 CCCAGGACGAGGCCCCCTGCTGCCACGCAACCCCTTCCCTATAGCAGACTGCCT 1023  
Db 965 CCCAGGACGAGGCCCCCTGCTGCCACGCAACCCCTTCCCTATAGCAGACTGCCT 1024  
Qy 1024 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCAACCTCTGGAGTCTTCTCTG 1083  
Db 1025 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCAACCTCTGGAGTCTTCTCTG 1084  
Qy 1084 GGCAAACTGGCTGTGTAGACAGCAGAAACCTTGGCTGACCCACACCTCAGCAAACTG 1143  
Db 1085 GGCAAACTGGCTGTGTAGACAGCAGAAACCTTGGCTGACCCACACCTGAGCAAACTG 1144  
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGCCAGCTGCTGTCT 1203  
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTGGAGCCAGCTGCTGTCT 1204  
Qy 1204 CGTGGTCCCATGAGAAGTTTCTGTGAGAGAACTAGTGTGATGCTCTGCAAGGAGC 1263  
Db 1205 CGTGGTCCCATGAGAAGTTTCTGTGAGAGAACTAGTGTGATGCTCTGCAAGGAGC 1264  
Qy 1264 GTGAGCTCAAGCCAGGCCCCAGCTGACAGCTGGCCAGAGCTGGGAGCAGCGGGC 1323  
Db 1265 GTGAGCTCAAGCCAGGCCCCAGCTGACAGCTGGCCAGAGCTGGGAGCAGCGGGC 1324  
Qy 1324 TCCAGATCCCGGAGGCCAGCCCCAAACTGAGGACCAACAGAGGTGTCCTCTCACTGAG 1383  
Db 1325 TCCAGATCCCGGAGGCCAGCCCCAAACTGAGGACCAACAGAGGTGTCCTCTCACTGAG 1384  
Qy 1384 AAATCAAGCCAGTGGATTTAGTATCCGAGAGAGTTCCTGGGCCAGCCACAGCTC 1443  
Db 1385 AAATCAAGCCAGTGGATTTAGTATCCGAGAGAGTTCCTGGGCCAGCCACAGCTC 1444  
Qy 1444 CGCTGGGCGAGGCTCTCGGAGAGTGCACAGGATGGAGGACAGCAGACTGGCTTC 1503  
Db 1445 CGCTGGGCGAGGCTCTCGGAGAGTGCACAGGATGGAGGACAGCAGACTGGCTTC 1504  
Qy 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAAGTATTTGGGCGAGGAGCTGATGGCATGT 1563  
Db 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAAGTATTTGGGCGAGGAGCTGATGGCATGT 1564  
Qy 1564 GCAGGATGACCTCAACCCAGAAATTTGCTTTGATGAGAGTGTGAGAGAGGCGCTTGG 1623  
Db 1565 GCAGGATGACCTCAACCCAGAAATTTGCTTTGATGAGAGTGTGAGAGAGGCGCTTGG 1624  
Qy 1624 GTCAACATTTCAAGGAGCTCTGGAGAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1683  
Db 1625 GTCAACATTTCAAGGAGCTCTGGAGAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1684  
Qy 1684 GGCTGTCTCCAGAGGACCGGCGCTGTACTTACCTGGGCGAGGCGCTGGAGGCTGTGAA 1743  
Db 1685 GGCTGTCTCCAGAGGACCGGCGCTGTACTTACCTGGGCGAGGCGCTGGAGGCTGTGAA 1744  
Qy 1744 TACCTCACTCAAGAGGATTTGATGATGGGAGCTCAAGAGTGTGACAGCTGCTCTCTGTC 1803  
Db 1745 TACCTCACTCAAGAGGATTTGATGATGGGAGCTCAAGAGTGTGACAGCTGCTCTCTGTC 1804

Qy 1804 AGCGATGGGAGCCACGACGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTTCAACCTGAT 1863  
Db 1805 AGCGATGGGAGCCACGACGCCCTCTGTGACTTTTGGCCATGCTGTGTCTTTCAACCTGAT 1864  
Qy 1864 GGCTTGGGAAAGTCTTTGCTCAAGGGGACTATCATCTCTGGCAGAGAGCCACATGGCT 1923  
Db 1865 GGCTTGGGAAAGTCTTTGCTCAAGGGGACTATCATCTCTGGCAGAGAGCCACATGGCT 1924  
Qy 1924 CCGGAGTGTGTGGGACGAGCTGCAGCCCAAGGTGAGATGCTGAGCAGCTGTGT 1983  
Db 1925 CCGGAGTGTGTGGGACGAGCTGCAGCCCAAGGTGAGATGCTGAGCAGCTGTGT 1984  
Qy 1984 ATGATGTGCACATGCTCAACGGTGCACCCCTTGGACTCAGTTCTTCCAGAGGCGCTC 2043  
Db 1985 ATGATGTGCACATGCTCAACGGTGCACCCCTTGGACTCAGTTCTTCCAGAGGCGCTC 2044  
Qy 2044 TGCCTCAAGATTGCGACGAGCTTCCGCTGTGTGAGGAGATCCACCTCTCTGCGCCCT 2103  
Db 2045 TGCCTCAAGATTGCGACGAGCTTCCGCTGTGTGAGGAGATCCACCTCTCTGCGCCCT 2104  
Qy 2104 CTACAGCCAGGCGCATCCAGAGGGCTGAGGAGAGCCCATCCAGCTGTCTGCA 2163  
Db 2105 CTACAGCCAGGCGCATCCAGAGGGCTGAGGAGAGCCCATCCAGCTGTCTGCA 2164  
Qy 2164 GCGGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2223  
Db 2165 GCGGAGCTGGGAGGAAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2224  
Qy 2224 TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAATCAAGGCAATTAACACAG 2283  
Db 2225 TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCAAAATCAAGGCAATTAACACAG 2284  
Qy 2284 ACCCTCATGCCAGCGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCCGGCCAGCTGAG 2343  
Db 2285 ACCCTCATGCCAGCGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCCGGCCAGCTGAG 2344  
Qy 2344 GAGCAACAGCGAGAGCCCTTAAGCTCAGCTCTCTCCACAGAGCCCCAGAGGCA 2403  
Db 2345 GAGCAACAGCGAGAGCCCTTAAGCTCAGCTCTCTCCACAGAGCCCCAGAGGCA 2404  
Qy 2404 AACAGTCTCTCCCTTGAATTTGAGCAAGGAGTCTGGGATGTGGAAACCTTACCT 2463  
Db 2405 AACAGTCTCTCCCTTGAATTTGAGCAAGGAGTCTGGGATGTGGAAACCTTACCT 2464  
Qy 2464 CTGTCTCTCTGGAGCCAGCCCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCTCTCTGGAGCCAGCCCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524  
Qy 2524 GTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAAATTTATTTCTCAACAGCTGTCCCAG 2583  
Db 2525 GTCCCGGAGCAGGAACTGCAGCAGCTGGAATAGAAATTTATTTCTCAACAGCTGTCCCAG 2584  
Qy 2584 CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTCCCTCAGCATCGACAGCTCTCC 2643  
Db 2585 CCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTCCCTCAGCATCGACAGCTCTCC 2644  
Qy 2644 CTGTGAGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGAGCACCCCTG 2703  
Db 2645 CTGTGAGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGAGCACCCCTG 2704  
Qy 2704 AGCTCAGGCGTACACTCTCGAGCAGCGGAGCTCGAAGCTCAGCTGGAACATG 2763  
Db 2705 AGCTCAGGCGTACACTCTCGAGCAGCGGAGCTCGAAGCTCAGCTGGAACATG 2764  
Qy 2764 GTGTGCGCGCGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAA 2823  
Db 2765 GTGTGCGCGCGGCGGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCCAA 2824  
Qy 2824 ATACAGTCTCTTAATGAGTGAACCTTGCAATCCGGAGTTCCACCGGGTCAAGTGGGA 2883  
Db 2825 ATACAGTCTCTTAATGAGTGAACCTTGCAATCCGGAGTTCCACCGGGTCAAGTGGGA 2884

2884 QY GACATCGCCATCTGGGATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAA 2943  
2885 Db GACATCGCCATCTGGGATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAA 2944  
2944 QY GACGGGACGCTGTTGCTAGCAGATGAGGTCAGAGCTCGGCGATCGACCTGAGTGC 3003  
2945 Db GACGGGACGCTGTTGCTAGCAGATGAGGTCAGAGCTCGGCGATCGACCTGAGTGC 3004  
3004 QY ACACCTGGCCCTGATGGAGCTTCGCCCTGGAGCTCGAGGCTCAAGCATGGCCAGCTGGAG 3063  
3005 Db ACACCTGGCCCTGATGGAGCTTCGCCCTGGAGCTCGAGGCTCAAGCATGGCCAGCTGGAG 3064  
3064 QY AACAGGCCCTAACCTGCGCTCCACCGCGGCTCCACATGCGGAAAGCAGCCTTCCTG 3123  
3065 Db AACAGGCCCTAACCTGCGCTCCACCGCGGCTCCACATGCGGAAAGCAGCCTTCCTG 3124  
3124 QY CTCGGTGCAGATGCTGCGCTGAAAACACAGGCTCAGCGGTTCCAGGGGATTCAGGCC 3183  
3125 Db CTCGGTGCAGATGCTGCGCTGAAAACACAGGCTCAGCGGTTCCAGGGGATTCAGGCC 3184  
3184 QY CCCCGGCTCACAGTGGGAACAGGCGCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCC 3243  
3185 Db CCCCGGCTCACAGTGGGAACAGGCGCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCC 3244  
3244 QY TCCCAGGATTTTACACCTGAGCCCTGCCCCACCCCTGCTGAAAAACATCCCGCCAGTGAA 3303  
3245 Db TCCCAGGATTTTACACCTGAGCCCTGCCCCACCCCTGCTGAAAAACATCCCGCCAGTGAA 3304  
3304 QY GAGACAGAGGAGATGGCAGAGTTACTTGGGGAAAAACAAACAGGGATCTTTTCTGCC 3363  
3305 Db GAGACAGAGGAGATGGCAGAGTTACTTGGGGAAAAACAAACAGGGATCTTTTCTGCC 3364  
3364 QY CTGCTCCAGTGCAGTGTGCGCTGACCGCTTGGATCAGTGACCATTTGTTGCGACAGCAGG 3423  
3365 Db CTGCTCCAGTGCAGTGTGCGCTGACCGCTTGGATCAGTGACCATTTGTTGCGACAGCAGG 3424  
3424 QY GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACCCTC 3483  
3425 Db GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACCCTC 3484  
3484 QY CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTTCAAGTGTGAGAACCA 3543  
3485 Db CAGGCTGCTGAGAGTGTCAAGTGTGTAAGGGGCCAAACTCAGGTTTCAAGTGTGAGAACCA 3544  
3544 QY GGTACAGAGTATGCCCGCCCTAGTGTAAAGGGGCCCTCTAAACCCCTTGCCTGCGCCTC 3603  
3545 Db GGTACAGAGTATGCCCGCCCTAGTGTAAAGGGGCCCTCTAAACCCCTTGCCTGCGCCTC 3604  
3604 QY ACCTGGCCAGCTCACCCCTTTTGGGTGTAAGGGGAAAAAGAAATGCCTGAGCCCTGGGAAGGT 3663  
3605 Db ACCTGGCCAGCTCACCCCTTTTGGGTGTAAGGGGAAAAAGAAATGCCTGAGCCCTGGGAAGGT 3664  
3664 QY CCTGGTAGAATACACACATTTTCAGGTTGTTGGAAACACAGGTCCTGAGTTGACCTCT 3723  
3665 Db CCTGGTAGAATACACACATTTTCAGGTTGTTGGAAACACACAGGTCCTGAGTTGACCTCT 3724  
3724 QY GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTCTCAGTCCCGCAGACATGT 3783  
3725 Db GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTCTCAGTCCCGCAGACATGT 3784  
3784 QY GCCCTTTGCTGCTGCTACCACTCTTCCAGAGCAGCAGGCGCCCGAGCCCTTCAGGC 3843  
3785 Db GCCCTTTGCTGCTGCTACCACTCTTCCAGAGCAGCAGGCGCCCGAGCCCTTCAGGC 3844  
3844 QY CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGAT 3903  
3845 Db CCAGCACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTTGAT 3904  
3904 QY GCAGGATGCTGTGACAGAAACAGTCTGGATGACATGATCAGTGTCAAGGAAGCAG 3963  
3905 Db GCAGGATGCTGTGACAGAAACAGTCTGGATGACATGATCAGTGTCAAGGAAGCAG 3964  
3964 QY CAGAGAGACGTCCGGCGCCCCCAGCCCCACTATCAGTGTCCAGCGTGTGTTCCCCCAG 4023

3965 Db CAGAGAGAGCGTCCGGCGCCCCCAGCCCACTATCAGTGTCCAGCGTCTGTGTTCCCCAG 4024  
4024 QY AGCAGAGCTCAGCATCACACTCACCCTGCTGCCCTGCTGCCCTGCGCCAGAGGCTACTGC 4083  
4025 Db AGCAGAGCTCAGCATCACACTCACCCTGCTGCCCTGCTGCCCTGCGCCAGAGGCTACTGC 4084  
4084 QY CGACGGCACTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGTGCCTCTGGCAGGGC 4143  
4085 Db CGACGGCACTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGTGCCTCTGGCAGGGC 4144  
4144 QY AGGGCAGGGCAGTGCACACTGTAGGAGCATACAGCCAGGAGATGGGTTGAAGGACACA 4203  
4145 Db AGGGCAGGGCAGTGCACACTGTAGGAGCATACAGCCAGGAGATGGGTTGAAGGACACA 4204  
4204 QY GTCTTTGAGCTCTGCACATGCACTGACCTCAAAACCTTCTCCAGATTTCTTAAGAATA 4263  
4205 Db GTCTTTGAGCTCTGCACATGCACTGACCTCAAAACCTTCTCCAGATTTCTTAAGAATA 4264  
4264 QY GCACCCCTTCCCACTTGGCCCGCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC 4323  
4265 Db GCACCCCTTCCCACTTGGCCCGCAGCTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC 4324  
4324 QY ACCTAGCATTTAAATCAGCTGTGAATCTGCAGGGGTTGCTGCTAGCTCAACCTCCTGGG 4383  
4325 Db ACCTAGCATTTAAATCAGCTGTGAATCTGCAGGGGTTGCTGCTAGCTCAACCTCCTGGG 4384  
4384 QY CGAGGGACGCGGAGACTCCGTGGGAGAGCTCATTCCACATCTTGCACAGACAGCCTT 4443  
4385 Db CGAGGGACGCGGAGACTCCGTGGGAGAGCTCATTCCACATCTTGCACAGACAGCCTT 4444  
4444 QY TGTCCAGCTGCCACATTCAGTTCAGCTGCTCCCGGGGAGAGAGCCCGGCCCCCGCAGCAC 4503  
4445 Db TGTCCAGCTGCCACATTCAGTTCAGCTGCTCCCGGGGAGAGAGCCCGGCCCCCGCAGCAC 4504  
4504 QY ATAAAGAACTCGAGCCTTGGTACTGACAGTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4563  
4505 Db ATAAAGAACTCGAGCCTTGGTACTGACAGTCTGGGTTGTAGAGAACTCTTTGTAAGCAA 4564  
4564 QY TAAAGTTTGGGTTGATGACAAATGTTAAAAA 4595  
4565 Db TAAAGTTTGGGTTGATGACAAATGTTAAAAA 4596

RESULT 4

US-10-283-975A-294  
; Sequence 294, Application US/10283975A  
; Publication No. US20040110792A1  
; GENERAL INFORMATION:  
; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
; FILE REFERENCE: CDS 293 PCT  
; CURRENT APPLICATION NUMBER: US/10/283,975A  
; PRIOR FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: 60/340,938  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/338,997  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,081  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/341,012  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-283-975A-294

Query Match 99.6%; Score 4578.4; DB 19; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;



QY 5 GGGGACGTGTCGCTGTGGAACTGTAGCTTGA-AGTGGACTCTGTACCAATTGAGG 63  
Db |||||  
QY 5 GGGGACGTGTCGCTGTGGAACTGTAGCTTGA-AGTGGACTCTGTACCAATTGAGG 64  
Db |||||  
QY 64 ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGAGACCTTTGGCC 123  
Db |||||  
QY 65 ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGAGACCTTTGGCC 124  
Db |||||  
QY 124 CCTGCCCTTTCGCCCAACCAAGGCTGACCTGTGTCTCCAGGCTTGGGATCTAAAGT 183  
Db |||||  
QY 125 CCTGCCCTTTCGCCCAACCAAGGCTGACCTGTGTCTCCAGGCTTGGGATCTAAAGT 184  
Db |||||  
QY 184 GACCTGCTCTGTGTGTGTGTCTCTCAGGATGAGCAGACAGCCTGGGAGATGGCAGTATG 243  
Db |||||  
QY 185 GACCTGCTCTGTGTGTGTGTCTCTCAGGATGAGCAGACAGCCTGGGAGATGGCAGTATG 244  
Db |||||  
QY 244 GAAATGGCTGCCAGAGTGGCTTGGCTCAGAGTGGGAGAGAAAGAACTCCCAAG 303  
Db |||||  
QY 245 GAAATGGCTGCCAGAGTGGCTTGGCTCAGAGTGGGAGAGAAAGAACTCCCAAG 304  
Db |||||  
QY 304 CCAAGGAGAGAGACGCGGCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGGAGCC 363  
Db |||||  
QY 305 CCAAGGAGAGAGACGCGGCCACTGGGGAAGAAACAGAGCTCCGTCTACAGCTTGGAGCC 364  
Db |||||  
QY 364 GTGAGAGAGACCTGTGTCTGTGGGAAAGTGGGAGATCTTGAATGACGTGATTTACCAAG 423  
Db |||||  
QY 365 GTGAGAGAGACCTGTGTCTGTGGGAAAGTGGGAGATCTTGAATGACGTGATTTACCAAG 424  
Db |||||  
QY 424 GGACACGCCAGGAGAGCTCCGAGCAGGCGCAGCTGCCATCTATCATCGCCAGGCT 483  
Db |||||  
QY 425 GGACACGCCAGGAGAGCTCCGAGCAGGCGCAGCTGCCATCTATCATCGCCAGGCT 484  
Db |||||  
QY 484 GAGTGTGAGATAGCCAAAGAGTTGAGCCCAACCTTTTCAGAAACGATTTTCATCGCTGGG 543  
Db |||||  
QY 485 GAGTGTGAGATAGCCAAAGAGTTGAGCCCAACCTTTTCAGAAACGATTTTCATCGCTGGG 544  
Db |||||  
QY 544 TCCAAAACAGTACGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT 603  
Db |||||  
QY 545 TCCAAAACAGTACGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT 604  
Db |||||  
QY 604 ACAGAGGCAAAATGGCCCGTGTGTGTGGAGGGAAGCGTCGACGAAAGCCGGAAG 663  
Db |||||  
QY 605 ACAGAGGCAAAATGGCCCGTGTGTGTGGAGGGAAGCGTCGACGAAAGCCGGAAG 664  
Db |||||  
QY 664 AAACGGAAGAAAGAGTCAAAAGTCCCTGGCTCATGACGAGTGGCTTGGCCCAACCC 723  
Db |||||  
QY 665 AAACGGAAGAAAGAGTCAAAAGTCCCTGGCTCATGACGAGTGGCTTGGCCCAACCC 724  
Db |||||  
QY 724 CTCCCCAGGACCCCTGAGCAGGAGTGCACCAATCCAGTGCAGGAGTGAATCTCCA 783  
Db |||||  
QY 725 CTCCCCAGGACCCCTGAGCAGGAGTGCACCAATCCAGTGCAGGAGTGAATCTCCA 784  
Db |||||  
QY 784 CTCGGGCCCCATATGTTAGAAACACCCCGAGTTTACCAAGCCTCTGAAGGAACACAGG 843  
Db |||||  
QY 785 CTCGGGCCCCATATGTTAGAAACACCCCGAGTTTACCAAGCCTCTGAAGGAACACAGG 844  
Db |||||  
QY 844 CTTGGGCAACTCTGTTTTAAGCAGCTTGGGAGGCGCTAGGCGCGCTCTGCCCTCGATCA 903  
Db |||||  
QY 845 CTTGGGCAACTCTGTTTTAAGCAGCTTGGGAGGCGCTAGGCGCGCTCTGCCCTCGATCA 904  
Db |||||  
QY 904 GAACTCCACAACTGATCAGCCCTTGCATATGTTGAACACACAGTGTGGAAACTGCACAC 963  
Db |||||  
QY 905 GAACTCCACAACTGATCAGCCCTTGCATATGTTGAACACACAGTGTGGAAACTGCACAC 964  
Db |||||  
QY 964 CCCCAGAGCGAGGCCCCCTGCCCCCAGCAGCAACCCCTTCCCTATAGCAGACTGCCT 1023  
Db |||||  
QY 965 CCCCAGAGCGAGGCCCCCTGCCCCCAGCAGCAACCCCTTCCCTATAGCAGACTGCCT 1024  
Db |||||  
QY 1024 CATCCCTTCCCATTCACCCCTTCCAGCCCTGGAACTTCACTCTGGAGTCTTCTCTG 1083  
Db |||||  
QY 1025 CATCCCTTCCCATTCACCCCTTCCAGCCCTGGAACTTCACTCTGGAGTCTTCTCTG 1084  
Db |||||

QY 1084 GGCNAACCTGGCCTGTGTAGACGCCAGAAACCTTGGCTGTGACCCACACCTTGAGCAAACTG 1143  
Db |||||  
QY 1085 GGCNAACCTGGCCTGTGTAGACGCCAGAAACCTTGGCTGTGACCCACACCTTGAGCAAACTG 1144  
Db |||||  
QY 1144 GCCTGTGTAGACAGTCAAAAGCCCTGCTGGGCCACACCTGGAGCCAGCTGCTGTCT 1203  
Db |||||  
QY 1145 GCCTGTGTAGACAGTCAAAAGCCCTGCTGGGCCACACCTGGAGCCAGCTGCTGTCT 1204  
Db |||||  
QY 1204 CGTGGTCCCATGAGAAAGTTTCTGTGGAGGAATACTAGTGCATGTCTCTGAAAGGACG 1263  
Db |||||  
QY 1205 CGTGGTCCCATGAGAAAGTTTCTGTGGAGGAATACTAGTGCATGTCTCTGAAAGGACG 1264  
Db |||||  
QY 1264 GTGAGCTCAAGCAGGCCACAGCTCAGCAGCTGCGGAGCAAGACCTGGGAGCAGCGGGC 1323  
Db |||||  
QY 1265 GTGAGCTCAAGCAGGCCACAGCTCAGCAGCTGCGGAGCAAGACCTGGGAGCAGCGGGC 1324  
Db |||||  
QY 1324 TCCAGATCCCGGAGGCCAGCCCAAACTGAGGACAACTGAGGGTGTCTGCTCACTGAG 1383  
Db |||||  
QY 1325 TCCAGATCCCGGAGGCCAGCCCAAACTGAGGACAACTGAGGGTGTCTGCTCACTGAG 1384  
Db |||||  
QY 1384 AAACCTCAAGCAGTGGATTAAGTACCGAGAAAGTCCACTGGGCCACGCAACAGCTC 1443  
Db |||||  
QY 1385 AAACCTCAAGCAGTGGATTAAGTACCGAGAAAGTCCACTGGGCCACGCAACAGCTC 1444  
Db |||||  
QY 1444 CGCTGGGAGAGAGCTCCTTCGGAGAGGTGCAAGGATGAGGAGCAAGAGACACTGGCTTC 1503  
Db |||||  
QY 1445 CGCTGGGAGAGAGCTCCTTCGGAGAGGTGCAAGGATGAGGAGCAAGAGACACTGGCTTC 1504  
Db |||||  
QY 1504 CAGTGCCTCTCAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGGAGCTGATGGCATGT 1563  
Db |||||  
QY 1505 CAGTGCCTCTCAAAAGGTGCGCTGGAAGTATTTTCGGGAGAGGAGCTGATGGCATGT 1564  
Db |||||  
QY 1564 GCAGGATTGACCTCACCCAGAAATTTGCTCCTTTGTATGGAGCTGTGAGAAAGGCGCTTGG 1623  
Db |||||  
QY 1565 GCAGGATTGACCTCACCCAGAAATTTGCTCCTTTGTATGGAGCTGTGAGAAAGGCGCTTGG 1624  
Db |||||  
QY 1624 GTCAACATCTTTCAGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGCTGCTCAAGAGAGCAG 1683  
Db |||||  
QY 1625 GTCAACATCTTTCAGAGCTGCTGGAAGGTGGCTCCCTGGGCCAGCTGCTCAAGAGAGCAG 1684  
Db |||||  
QY 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTGGAA 1743  
Db |||||  
QY 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTGGAA 1744  
Db |||||  
QY 1744 TACTCCACTCAGAAAGATTTCTCATGGGAGAGTCAAAAGCTGACAACTGCTCTCTGTCC 1803  
Db |||||  
QY 1745 TACTCCACTCAGAAAGATTTCTCATGGGAGAGTCAAAAGCTGACAACTGCTCTCTGTCC 1804  
Db |||||  
QY 1804 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCCATGCTGTGTCTTTCAACCTGAT 1863  
Db |||||  
QY 1805 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCCATGCTGTGTCTTTCAACCTGAT 1864  
Db |||||  
QY 1864 GGCTTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCT 1923  
Db |||||  
QY 1865 GGCTTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCCACAGAGACCCACATGGCT 1924  
Db |||||  
QY 1924 CCGAGGTGTGTCTGGGAGAGCTGCGAGCCAAAGGTGATGTCTGGAGCAGCTGTGT 1983  
Db |||||  
QY 1925 CCGAGGTGTGTCTGGGAGAGCTGCGAGCCCAAGGTGATGTCTGGAGCAGCTGTGT 1984  
Db |||||  
QY 1984 ATGATGTGCATGTCTCAACGGCTGCACCCCTTGGACTCAGTTCTTCCGAGGCGCGCTC 2043  
Db |||||  
QY 1985 ATGATGTGCATGTCTCAACGGCTGCACCCCTTGGACTCAGTTCTTCCGAGGCGCGCTC 2044  
Db |||||  
QY 2044 TGCTTCAAGATTGGCAGGAGCTTCCGCTGTGAGGAGATCCACCCCTCTCTGGCGCCCT 2103  
Db |||||  
QY 2045 TGCTTCAAGATTGGCAGGAGCTTCCGCTGTGAGGAGATCCACCCCTCTCTGGCGCCCT 2104  
Db |||||  
QY 2104 CTCACAGCCAGGAGCTTCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTGCA 2163  
Db |||||  
QY 2105 CTCACAGCCAGGAGCTTCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGGCTGTCTGCA 2164  
Db |||||  
QY 2164 GCGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223  
Db |||||

Db	2165	 GCGGAGCTGGGAGGAAAGGTCAACCGGGCACTACGCAAGTGGGAGGTCTTGAGAGCCCT	2224
QY	2224	 TGGAGGGGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAATTAACCAACCAG	2283
Db	2225	 TGGAGGGGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAATTAACCAACCAG	2284
QY	2284	 ACCCCTCATGCCCCACCGCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAGCTGAG	2343
Db	2285	 ACCCCTCATGCCAGCCGAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCGGCCAGCTGAG	2344
QY	2344	 GAGACAAAGCAGAGCCCCCTAAGCTCCAGCTCTCTCTCCCAACAGAGCCCCCAGAGCCA	2403
Db	2345	 GAGACAAAGCAGAGCCCCCTAAGCTCCAGCTCTCTCTCCCAACAGAGCCCCCAGAGCCA	2404
QY	2404	 AACAAAGTCTCCTCCCTTGACTTTTTCAGCAAGAGGAGTCTGGGATGTGGGAACCCCTTACCT	2463
Db	2405	 AACAAAGTCTCCTCCCTTGACTTTTTCAGCAAGAGGAGTCTGGGATGTGGGAACCCCTTACCT	2464
QY	2464	 CTGTCTCTCCCTGGAGCCAGCCCTTCCAGAGAAACCCAGCTCACACAGAGCGGAAGCAACC	2523
Db	2465	 CTGTCTCTCCCTGGAGCCAGCCCTTCCAGAGAAACCCAGCTCACACAGAGCGGAAGCAACC	2524
QY	2524	 GTCCCGGACGAGGAACCTGCAGAGCTGGAATATTAATTCCTCAACAGCCTGTCTCCAG	2583
Db	2525	 GTCCCGGACGAGGAACCTGCAGAGCTGGAATATTAATTCCTCAACAGCCTGTCTCCAG	2584
QY	2584	 CCATTTTCTCTGGAGGAGCAGGAGCAAAATTCCTCTCGTCTCAGCATCGACGCTCTCC	2643
Db	2585	 CCATTTTCTCTGGAGGAGCAGGAGCAAAATTCCTCTCGTCTCAGCATCGACGCTCTCC	2644
QY	2644	 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAAAGCTCGCGGACACCCCTG	2703
Db	2645	 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAAAGCTCGCGGACACCCCTG	2704
QY	2704	 AGCTCAGGCGTACACTCCTGGAGCAGCCAGGCTCGAAGCTCGAAGCTCCAGCTGGAAATG	2763
Db	2705	 AGCTCAGGCGTACACTCCTGGAGCAGCCAGGCTCGAAGCTCGAAGCTCCAGCTGGAAATG	2764
QY	2764	 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTTCATGGTGTGAAGTCCAA	2823
Db	2765	 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTTCATGGTGTGAAGTCCAA	2824
QY	2824	 ATACAGTCTCTTAATGGTGAACACTGCACATCGGGAGTTTCCACCGGTCAAAGTGGGA	2883
Db	2825	 ATACAGTCTCTTAATGGTGAACACTGCACATCGGGAGTTTCCACCGGTCAAAGTGGGA	2884
QY	2884	 GACATCGCCACTGGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGCTTGGTCAACAAA	2943
Db	2885	 GACATCGCCACTGGCATCAGCAGCCAGATCCAGCTCAGCCTTCAGCTTGGTCAACAAA	2944
QY	2944	 GACGGGACGCTGTTCGTCAGCATATGGAGGTGCAGACTCGGGCATCGACCTGCGAGTGC	3003
Db	2945	 GACGGGACGCTGTTCGTCAGCATATGGAGGTGCAGACTCGGGCATCGACCTGCGAGTGC	3004
QY	3004	 ACACTGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGTCGAAGCATGCGCAGCTGGAG	3063
Db	3005	 ACACTGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGTCGAAGCATGCGCAGCTGGAG	3064
QY	3064	 AACAGGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGCGGAAGCAGCCTTCCCTG	3123
Db	3065	 AACAGGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGCGGAAGCAGCCTTCCCTG	3124
QY	3124	 TTCGTGTACGATGTGCTGAAAAACACAGGCTCAGCGTTTCCAGGGGATGTCAGACC	3183
Db	3125	 TTCGTGTACGATGTGCTGAAAAACACAGGCTCAGCGTTTCCAGGGGATGTCAGACC	3184
QY	3184	 CCCCGGCTCAGAGTGGGAACACAGGSCCTCGCAGCAGCAAGGTGGGGCAGCAGAAATGCC	3243
Db	3185	 CCCCGGCTCAGAGTGGGAACACAGGSCCTCGCAGCAGCAAGGTGGGGCAGCAGAAATGCC	3244
QY	3244	 TCCCAGGATTTTCAACCTTGAGCCCTGCCCCCACCTGTGTGAAAAAATATCCGCCACGTGAA	3303

Db	3245	TCCAGAGATTTCA	CCTGAGCCCTG	CCGCCA	CCTCTGCTG	AAAAAACAT	CGGCCACGTG	AA	3304		
Qy	3304	GAGACAAAGAGATG	GCAGGATTA	CTACCTGGGGAAA	CAAAAC	CAGGGATCT	TTTTCTGCC		3363		
Db	3305	GAGACAAAGAGATG	GCAGGATTA	CTACCTGGGGAAA	CAAAAC	CAGGGATCT	TTTTCTGCC		3364		
Qy	3364	CTTGCTCCAGT	CGAGTTGGCT	GACCCGCTTG	GAATCAGT	GACCAATTTG	TGTGGCAGACAGG		3423		
Db	3365	CTTGCTCCAGT	TCGAGTTGGCT	GACCCGCTTG	GAATCAGT	GACCAATTTG	TGTGGCAGACAGG		3424		
Qy	3424	GGAGACAGCTT	CCAGCTGGGT	CAGAAAGGGT	GGCGAGCCCT	TCGGGCCCT	TCACCCCTC		3483		
Db	3425	GGAGACAGCTT	CCAGCTGGGT	CAGAAAGGGT	GGCGAGCCCT	TCGGGCCCT	TCACCCCTC		3484		
Qy	3484	CAGGCTGCTG	GAGAGTGT	CAAGTGTGTA	AGGGCCCAAA	CTCAGGTTCA	GTGTCAGCAGAACCA		3543		
Db	3485	CAGGCTGCTG	GAGAGTGT	CAAGTGTGTA	AGGGCCCAAA	CTCAGGTTCA	GTGTCAGCAGAACCA		3544		
Qy	3544	GGTCAGCAGGTA	TGCCCGCTG	AGTGTAA	AGGGGGCCCT	TAAACCCCT	TGCTGGCCCTC		3603		
Db	3545	GGTCAGCAGGTA	TGCCCGCTG	AGTGTAA	AGGGGGCCCT	TAAACCCCT	TGCTGGCCCTC		3604		
Qy	3604	ACCTGGCCAGCT	CACCCCTTTT	GGGTGTAG	GGGGAAG	AATGCTC	CACCTCGGGAAGGT		3663		
Db	3605	ACCTGGCCAGCT	CACCCCTTTT	GGGTGTAG	GGGGAAG	AATGCTC	CACCTCGGGAAGGT		3664		
Qy	3664	CCCTGGTAGAAT	AACACACTTT	TTCAGGTTG	TGCAACAC	CAGGCTCT	CGTAGTTGACCTCT		3723		
Db	3665	CCCTGGTAGAAT	AACACACTTT	TTCAGGTTG	TGCAACAC	CAGGCTCT	CGTAGTTGACCTCT		3724		
Qy	3724	GGTTCCAGCA	AGGACAAAG	AGTGTGA	GTGAAGTGT	TTCTAGT	CCCCCAGACATGT		3783		
Db	3725	GGTTCCAGCA	AGGACAAAG	AGTGTGA	GTGAAGTGT	TTCTAGT	CCCCCAGACATGT		3784		
Qy	3784	GCCCTTTGCT	GTGGCTAC	CACTCTCC	CCAGCAGC	CAGGCCCC	CGCTTCAGGC		3843		
Db	3785	GCCCTTTGCT	GTGGCTAC	CACTCTCC	CCAGCAGC	CAGGCCCC	CGCTTCAGGC		3844		
Qy	3844	CCAGCACT	TGCCCA	GACATCG	CTGGCACT	CAGTTCCCT	CATCTCTAAAGGTGA	AGGGTGAT	3903		
Db	3845	CCAGCACT	TGCCCA	GACATCG	CTGGCACT	CAGTTCCCT	CATCTCTAAAGGTGA	AGGGTGAT	3904		
Qy	3904	CGAGGATATG	CTGACAGGA	CAAGTCTGT	GGATGG	ACATGAT	CAGTGTCTTA	AGGAAAGCAG	3963		
Db	3905	CGAGGATATG	CTGACAGGA	CAAGTCTGT	GGATGG	ACATGAT	CAGTGTCTTA	AGGAAAGCAG	3964		
Qy	3964	CAGAGAGAGCT	CCGGGGCC	CAGCCCC	ACTATC	AGTGTCC	AGCGTGTCTG	TTCCCCCAG	4023		
Db	3965	CAGAGAGAGCT	CCGGGGCC	CAGCCCC	ACTATC	AGTGTCC	AGCGTGTCTG	TTCCCCCAG	4024		
Qy	4024	AGCACAGCT	CAGCATCA	CACTGAC	ATCA	CCCTGCC	CTGCCCTTC	GGCAGAGGGTACTGC	4083		
Db	4025	AGCACAGCT	CAGCATCA	CACTGAC	ATCA	CCCTGCC	CTGCCCTTC	GGCAGAGGGTACTGC	4084		
Qy	4084	CGACGGCA	CTTTTGCA	CTGTGAT	GCCTCA	AAAGCA	CTTTTCA	TGGTGCCCTCT	GGCAGGGC	4143	
Db	4085	CGACGGCA	CTTTTGCA	CTGTGAT	GCCTCA	AAAGCA	CTTTTCA	TGGTGCCCTCT	GGCAGGGC	4144	
Qy	4144	AGGGCAGGGC	AGTGCAC	ACTGTAG	GAGCATAG	CACAGCC	AGGAGATGG	GGTCAAGGGACACA	4203		
Db	4145	AGGGCAGGGC	AGTGCAC	ACTGTAG	GAGCATAG	CACAGCC	AGGAGATGG	GGGTAAAGGGGACACA	4204		
Qy	4204	GTCTTTGAGCT	GTCCACATG	CATGTG	ACTCTCT	CAAAACCT	CTTCCAGATTT	CTCTAAAGAAATA	4263		
Db	4205	GTCTTTGAGCT	GTCCACATG	CATGTG	ACTCTCT	CAAAACCT	CTTCCAGATTT	CTCTAAAGAAATA	4264		
Qy	4264	GCACCCCTT	CCCCATTTGCC	CCAGCTTT	CTTCC	CAGGGGAG	ACTCTAG	GAGCTCTC	AGGACTC	4323	
Db	4265	GCACCCCTT	CCCCATTTGCC	CCAGCTTT	CTTCC	CAGGGGAG	ACTCTAG	GAGCTCTC	AGGACTC	4324	
Qy	4324	ACGTAGCA	TTAAATC	AGCTGTGA	ATCGTC	CAGGGGGT	GTCTGT	AGCTCTCA	ACCTCTCT	CGGG	4383
Db	4325	ACGTAGCA	TTAAATC	AGCTGTGA	ATCGTC	CAGGGGGT	GTCTGT	AGCTCTCA	ACCTCTCT	CGGG	4384





QY	1502	TCCAGTGGCGTGTCAA	AAAGTGGCCCTGGAAAGTATTTCCGGCAGAGGAGCTGATGGCAT	1561
DB	1553	TCCAGTGGCGTGTCAA	AAAGTGGCGGTGGAAAGTATTTCCGGCAGAGGAGCTGATGGCAT	1612
QY	1562	GTGCAGGATTGACCTCA	CCCAAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGGCTT	1621
DB	1613	GTGCAGGATTGACCTCA	CCCAAGAAATGTCCCTTTGTATGGAGCTGTGAGAGAGGGGCTT	1672
QY	1622	GGGTCAACATCTTTCAT	GAGAGCTGTGAGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGC	1681
DB	1673	GGGTCAACATCTTTCAT	GAGAGCTGTGAGAAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGC	1732
QY	1682	AGGCTGTCTCCACAGAG	ACCGGCCCTGTACTACTTGGGCCAGGCCCTGGAGGGTCTGG	1741
DB	1733	AGGCTGTCTCCACAGAG	ACCGGCCCTGTACTACTTGGGCCAGGCCCTGGAGGGTCTGG	1792
QY	1742	AATACCTCCACTCAAG	AAGGATTTCTGATGGGGAGCTCAAAAGCTGACAAAGTGTCTCTGT	1801
DB	1793	AATACCTCCACTCAAG	AAGGATTTCTGATGGGGAGCTCAAAAGCTGACAAAGTGTCTCTGT	1852
QY	1802	CCAGGGATGGAGGCGCA	CGCAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTG	1861
DB	1853	CCAGCGATGGAGGCGCA	CGCAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTG	1912
QY	1862	ATGGCTCTGGAAAGTCC	TTTGCTCACAGGGGACTACATCCCTGGGCACAGAGACCCACATGG	1921
DB	1913	ATGGCTCTGGAAAGTCC	TTTGCTCACAGGGGACTACATCCCTGGGCACAGAGACCCACATGG	1972
QY	1922	CTCCGGAGTGGTGTCTG	GGCAGGAGCTGCCAGCGCAAGGTGGATGTCTGGAGCAGTGTCT	1981
DB	1973	CTCCGGAGTGGTGTCTG	GGCAGGAGCTGCCAGCGCAAGGTGGATGTCTGGAGCAGTGTCT	2032
QY	1982	GTATGATGCTGCACATG	CTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGC	2041
DB	2033	GTATGATGCTGCACATG	CTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGC	2092
QY	2042	TCTGCTCTCAAGATTGC	CCAGCAGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGCCCC	2101
DB	2093	TCTGCTCTCAAGATTGC	CCAGCAGAGCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGCCCC	2152
QY	2102	CTCTCACAGCCCGAGCC	ATCCAGAGGGGTGAGGAAAGGCCCATCCACGGGTGTCTG	2161
DB	2153	CTCTCACAGCCCGAGCC	ATCCAGAGGGGTGAGGAAAGGCCCATCCACGGGTGTCTG	2212
QY	2162	CAGCGGAGCTGGAGGGA	AGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTCAAGAGCC	2221
DB	2213	CAGCGGAGCTGGAGGGA	AGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTCAAGAGCC	2272
QY	2222	CTTGGAGGGGAGATATA	AAAGAACACAGACATCCACCGCCAAATCAAGCCAAATTCACACC	2281
DB	2273	CTTGGAGGGGAGATATA	AAAGAACACAGACATCCACCGCCAAATCAAGCCAAATTCACACC	2332
QY	2282	AGACCTCCATGCCACG	CCGAGAGCTTTTCGCAAGGGCCCAAGGGCCCGGCCACGCTG	2341
DB	2333	AGACCTCCATGCCACG	CCGAGAGAGCTTTTCGCAAGGGCCCAAGGGCCCGGCCACGCTG	2392
QY	2342	AGGAGACAAAGGAGAG	CGCCCTTAAGCTTCAGGCTCTCTCCACACAGAGCCCCCAAGAGC	2401
DB	2393	AGGAGACAAAGGAGAG	CGCCCTTAAGCTTCAGGCTCTCTCCACACAGAGCCCCCAAGAGC	2452
QY	2402	CAAAACAGTCTCCCTTG	ACTTTTGACCAAGGAGGTCTGGGATGTGGGAACCTTAC	2461
DB	2453	CAAAACAGTCTCCCTTG	ACTTTTGACCAAGGAGGTCTGGGATGTGGGAACCTTAC	2512
QY	2462	CTCTGTCTCTCCCTTG	AGCCAGCCCTGCGAAGAACCCAGCTCACACAGAGCGGAAAGCAA	2521
DB	2513	CTCTGTCTCTCCCTTG	AGCCAGCCCTGCGAAGAACCCAGCTCACACAGAGCGGAAAGCAA	2572
QY	2522	CCGTCCCGGAGCAGGA	ACTGACAGAGCTGGAAATAGAAATTTCTCTCAAAGCTGTCCC	2581
DB	2573	CCGTCCCGGAGCAGGA	ACTGACAGAGCTGGAAATAGAAATTTCTCTCAAAGCTGTCCC	2632
QY	2582	AGCCATTTTCTCTGG	AGAGCAGGAGCAAAATTTCTCTGTGCTCTCAGCATCGACAGCTCT	2641

[illegible]



Db 781 AGCAGATGCTCATCCCTTCCATTCCACCCCTCTCCAGCCCTCGGAAACCTCACCCCTCTG 840  
Qy 1072 GAGTCTTCTCTGGGCAAACTGSCCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACAC 1131  
Db 841 GAGTCTTCTCTGGGCAAACTGSCCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACAC 900  
Qy 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTCGCTGGGCCACACCTCGAGGCC 1191  
Db 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTCGCTGGGCCACACCTCGAGGCC 960  
Qy 1192 AGCTGCCCTGTCTGTGGTCCCATGAGAAATTTCTGTGGAGAAATCACTAGTGCATGCT 1251  
Db 961 AGCTGCCCTGTCTGTGGTGGCCCATGAGAAATTTCTGTGGAGAAATCACTAGTGCATGCT 1020  
Qy 1252 CTGCAAGGCAAGGTGAGTCAAGCCAGGCCACAGCCTTGACAGCCTGGCCCAAGACCTGG 1311  
Db 1021 CTGCAAGGCAAGGTGAGTCAAGCCAGGCCACAGCCTTGACAGCCTGGCCCAAGACCTGG 1080  
Qy 1312 GCAGCACGGGGCTCCAGATCCCGGGAGGCCAGCCCCAAATCTGAGGACAAACGAGGGTGTCT 1371  
Db 1081 GCAGCACGGGGCTCTAGATCCCGGGAGGCCAGCCCCAAATCTGAGGACAAACGAGGGTGTCT 1140  
Qy 1372 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATACCGAGAAAGATCCACTGGGCC 1431  
Db 1141 CTGCTCACTGAGAAACTCAAGCCAGTGGATTATGATACCGAGAAAGATCCACTGGGCC 1200  
Qy 1432 AGGCACACGCTCCGCTGGGAGAGGCTCTCTCGAGAGGTGCACAGATGGAGACAAAG 1491  
Db 1201 AGGCACACGCTCCGCTGGGAGAGGCTCTCTCGAGAGGTGCACAGATGGAGACAAAG 1260  
Qy 1492 CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCCCTGGAAGTATTTGGGACAGGAG 1551  
Db 1261 CAGACTGGCTTCCAGTGGCTGTCAAAAAGGTGGCCCTGGAAGTATTTGGGACAGGAG 1320  
Qy 1552 CTGATGGCATGTGCAGGATTTGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGA 1611  
Db 1321 CTGATGGCATGTGCAGGATTTGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGA 1380  
Qy 1612 GAAAGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671  
Db 1381 GAAAGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440  
Qy 1672 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1731  
Db 1441 GTCAAGGAGCAGGCTGTCTCCAGAGACCGGGCCCTGTACTACTGGGCCAGGCCCTG 1500  
Qy 1732 GAGGCTCTGGAATACCTCACTCACGAAAGGATTTGATGGGGAGCTCAAAAGCTGACAAC 1791  
Db 1501 GAGGCTCTGGAATACCTCACTCACGAAAGGATTTGATGGGGAGCTCAAAAGCTGACAAC 1560  
Qy 1792 GTGCTCTGTCCAGGATGGAGCCACCGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1851  
Db 1561 GTGCTCTGTCCAGGATGGAGCCACCGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT 1620  
Qy 1852 CTTCAACTGTATGGCTGGGAAGTCTTGTCTCACAGGGACTACATCCCTGGGCACAGAG 1911  
Db 1621 CTTCAACTGTATGGCTGGGAAGTCTTGTCTCACAGGGACTACATCCCTGGGCACAGAG 1680  
Qy 1912 ACCCAATGGCTCCGAGGTGTGTGGGCAGGAGCTGCGACGCCAAGGTGATGCTGTGG 1971  
Db 1681 ACCCAATGGCTCCGAGGTGTGTGGGCAGGAGCTGCGACGCCAAGGTGATGCTGTGG 1740  
Qy 1972 AGCAGCTGTGTATGATGTCACATGCTCAAGGCTGGCACCCTCGACTCAGTCTTCTTC 2031  
Db 1741 AGCAGCTGTGTATGATGTCACATGCTCAAGGCTGGCACCCTCGACTCAGTCTTCTTC 1800  
Qy 2032 CGAGGGCCGCTCTGCTCAAGATTGCGACGAGCTCCGCTGTGAGGGAGATCCACCC 2091  
Db 1801 CGAGGGCCGCTCTGCTCAAGATTGCGACGAGCTCCGCTGTGAGGGAGATCCACCC 1860  
Qy 2092 TCCTCGGCCCTCTCACAGCCAGGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCAC 2151  
Db 1861 TCCTCGGCCCTCTCACAGCCAGGCCATCCAAGAGGGGCTGAGGAAAGAGCCCATCCAC 1920

Qy 2152 CGCGTGTCTGCAGCGGAGCTCGGAGGGAAGGTGAACCGGGCACTACACGAAGTGGAGGT 2211  
Db 1921 CGCGTGTCTGCAGCGGAGCTCGGAGGGAAGGTGAACCGGGCACTACACGAAGTGGAGGT 1980  
Qy 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCCAAATCAAGCC 2271  
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAGAAACCAAGACATCCACCCGCCAAATCAAGCC 2040  
Qy 2272 AATTACCAACAGACCTTCATGCCAGCCGAGAGAGCTTTTGCACGAAGGGCCCCAGGGCCC 2331  
Db 2041 AATTACCAACAGACCTTCATGCCAGCCGAGAGAGCTTTTGCACGAAGGGCCCCAGGGCCC 2100  
Qy 2332 CGGCCAGCTGAGGAGAGCAACAGGAGAGCCCTTAAGCTCCAGCCTCCTCTCCACACAGAG 2391  
Db 2101 CGGCCAGCTGAGGAGAGCAACAGGAGAGCCCTTAAGCTCCAGCCTCCTCTCCACACAGAG 2160  
Qy 2392 CCCCAGAGCCAAACAAGTCTCTCCCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGG 2451  
Db 2161 CCCCAGAGCCAAACAAGTCTCTCCCTTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGG 2220  
Qy 2452 GAAACCTTTACCTCTGTCTCCTTGAGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAG 2511  
Db 2221 GAAACCTTTACCTCTGTCTCCTTGAGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAG 2280  
Qy 2512 CGGAAAGCAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAATATTCTCTCAAC 2571  
Db 2281 CGGAAAGCAACCGTCCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAATATTCTCTCAAC 2340  
Qy 2572 AGCTGTCTCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGGCTCAGCATC 2631  
Db 2341 AGCTGTCTCCAGCCATTTTCTCTGAGGAGCAGGAGCAAAATTTCTCTCGTGGCTCAGCATC 2400  
Qy 2632 GACAGCTCTCTCTGTCTGGATGACAGTGAAGAAACCCATCAAAGGCTCTCTCAAAGCTCG 2691  
Db 2401 GACAGCTCTCTCTGTCTGGATGACAGTGAAGAAACCCATCAAAGGCTCTCTCAAAGCTCG 2460  
Qy 2692 CGGGACACCTTGAGCTCAGGCGTACCTCTGGAGCAGCCAGGCGGAGCTCGAGCTCC 2751  
Db 2461 CGGGACACCTTGAGCTCAGGCGTACCTCTGGAGCAGCCAGGCGGAGCTCGAGCTCC 2520  
Qy 2752 AGCTGGAACTGTGTCTGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2811  
Db 2521 AGCTGGAACTGTGTCTGGCCCGGGCGGCCACCGACACCCCAAGCTATTTCAATGGT 2580  
Qy 2812 GTGAAGTCCAAATACAGTCTCTTAATGGTGAAACACTGTGCACATCCGGGAGTTCACCCGG 2871  
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGGTGAAACACTGTGCACATCCGGGAGTTCACCCGG 2640  
Qy 2872 GTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGC 2931  
Db 2641 GTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCTTCAGC 2700  
Qy 2932 TTGGTCAACAAAGACGGGACGCTTTCGCTACGACATGGAGGTGCCAGACTCGGGCATC 2991  
Db 2701 TTGGTCAACAAAGACGGGACGCTTTCGCTACGACATGGAGGTGCCAGACTCGGGCATC 2760  
Qy 2992 GACTGTAGTGCACATGTGGCCCTGTGATGGCAGCTTTCGCTGGAGCTGGAGGTCAAGCAT 3051  
Db 2761 GACTGTAGTGCACATGTGGCCCTGTGATGGCAGCTTTCGCTGGAGCTGGAGGTCAAGCAT 2820  
Qy 3052 GGCCAGCTGGAGAACAGGCCCTTAA 3075  
Db 2821 GGCCAGCTGGAGAACAGGCCCTTAA 2844

## RESULT 7

US-10-087-192-884  
; Sequence 884, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.

## ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; FILE OF INVENTION: CANCER

; FILE REFERENCE: 52945200122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: RastSeq for Windows Version 4.0

; SEQ ID NO 884

; LENGTH: 2828

; TYPE: DNA

; ORGANISM: Mus musculus

US-10-087-192-884

Query Match 44.4%; Score 2042.6; DB 13; Length 2828;

Best Local Similarity 83.4%; Pred. No. 0;

Matches 2377; Conservative 0; Mismatches 444; Indels 28; Gaps 4;

Qy	232	ATGGCAGTGTGAAATGGCTGCCAGAGTGCCTGGCTCAGCAGTGGGGCAGCAGAAG	291
Db	1	ATGGCCGTGTGAAATGGCTGCCCGGCACTCTCTGGGTGAGCAGTGGGGCAGCAGAAG	60
Qy	292	GAACTCCCAAGCCAAAGGAGAGACGCGCCACTGCTGGGGAAGAAACAGAGTCCGTCTAC	351
Db	61	GAGCTTCCCAAGCCAAAGGAGAGACAGTCACTGGGGAAGAAAGCAGAGTGCATCTTC	120
Qy	352	AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTGCGGAAGATGGGAGATCCTGAATGAC	411
Db	121	AAGCTTGAGGCGGTGGAGAGAGCCCTGTGTTCTGCGGAAGATGGGAGATCCTAAACGAC	180
Qy	412	GTGATTACCAAGGSCACAGCCCAAGAGGCTCCGAGCAGGCGCAGCTGCATCTCTATC	471
Db	181	GTGATTACCAAGGSCACAGCCCAAGAGGCTCTGAGGAGGACCACCGGCCAATCTCCATC	240
Qy	472	ATCCGCCAGGCTGAGTGTGAGAAATAGCCAAAGAGTTACGCCCCACCTTTTCAGAAACGCAAT	531
Db	241	ATCCGCCAGGCTGAAATGTGAGAAATAGCCAAAGAGTTACGCCCCACCTTTTCAGAGCGCAAT	300
Qy	532	TTGATCGTGGTCCAAACAGTACAGCCAGTCCAGAGTCTTGATCAGATCCCAACAAAT	591
Db	301	TTGATCGGCGGTTCACAGCAGTACAGCCAGTCTGAGAGTCTCGATCAAAATCCCAACAAAT	360
Qy	592	GTGSCCATGCTACAGAGGCAAAATGSCCGTGTGTTGGAGGGAAGACGTCGCGAC	651
Db	361	GTGSCCATGCAACTGAGGCAAAATGSCCGTGTGTTGGAGGGAAGACGTCACGCGC	420
Qy	652	AAAGCCCGGAAGAAACGGAAGAGAGCTCAAGTCCCTGGCTCATGCGAGGAGTGGCC	711
Db	421	AAAGCCCGGAAGAAACGTAAGAGAGAGTCTGAAGTCACTGGCCAGGCGAGTGGCC	480
Qy	712	TTGGCCAAACCCCTCCAGAGACCCCTGAGCAGAGAGTGCACATCCCATGTCAGAG	771
Db	481	TTAGCCAAAGCCCTCCAGAGACCCCTGAGCAAGAGAGTGTACCATCCCAAGTACAGGAA	540
Qy	772	GATGAGTCTCACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCCTCTG	831
Db	541	GATGAGTCTCACTAGGCAACCTTATGCGCAGAAATGTCTCCCAAGTTTCAACAGCCCTCTG	600
Qy	832	AAGGAACACAGGCTTTGGGCACTCTGTTTAAAGCAGCTTGGCGAGGGGCTACGCGCGGCT	891
Db	601	GGGGGACACAGGCTTTGGGCACTCTGTTTAAAGAAACAGATGAAGGCTCGGACCGGTA	660
Qy	892	CTGCCTCGATCAGAACTCCAAATCTGATCAGCCCTTTGCAATGTCTGAACCAAGTGTGG	951
Db	661	CTGCCTCGAGCAACTCCAAATCTGATCAGCCCTTTGCAATGTCTAAACCAAGTGTGG	720
Qy	952	AAATGACACACCCAGGAGCGGAGGCCCTGCGCCCTGCCCCAGCGACCCCTTCCCTAT	1011
Db	721	AAATGACACACCCAGGAGCGGAGGCCCTGCGCCCTGCCCCAGCGACCCCTTCCCTAT	780

```
Db      1860  CCACCTTCTTGGGCAACCCCTCAAGCCAGGCCATCAAGAGGGGCTGAGGAAGAGCCC 1919
Qy      2146  ATCCACGGGTGCTGACGGAGCTGGAGGAGGTGAACCGGGCACTACAGCAAGTG 2205
Db      1920  GTCCACCGAGCATCTGCCATGGAGCTTCGGAGAAAGTGGCAAGCACTACAGGAAGTG 1979
Qy      2206  GGAGGTCTGAAGAGCCCTTGGAGGGGAGATATAAAGAACCAAGACATCCACGCCAAT 2265
Db      1980  GGAGGTCTGAAGCCCTTGGAAAGGAGATATAAAGAACCAAGACCTCCACCCCAAGAC 2039
Qy      2266  CAAAGCCAAATTAACACAGACCTTCCATGCTCCAGCCGAGAGACTTTTCGCAAGGGCCCA 2325
Db      2040  CAAAGCCACCTGCGACAGACCTTACCTACTCGCGGAGAGA----- 2080
Qy      2326  GGGCCCGGCGAGCTGAGAGACACAGCAGAGCCCTTAAGCTCAGGCTCTCTCTCCA 2385
Db      2081  --GAACCCACGCCCAAGGCCAACACAGACGGGGTCTCTGAGCCTCAGCCTCTCTACCG 2138
Qy      2386  CCAGAGCCCCCAGAGCAAAACAGTCTCTCTTGTGACTTTGAGCAAGGAGGCTGGG 2445
Db      2139  CAGAAACACAGAACCGAGCAAGGCCAGCCCTGAACTGAGCAAGGAGGAGTCTGGC 2198
Qy      2446  ATGTGGAAACCTTACTCTGTCTCTCTCTGAGCCAGCCCTTGCAGAAACCCCAAGCTCA 2505
Db      2199  ACATGGGAACCCCTGCTCTCTCTCTGAGCCAGCCCAAGCCACTGCGAAAGGCCCGAGCTTC 2258
Qy      2506  CCAGAGCGGAACCAACCGTCCGGAGCAGGAAGCTGACAGCAGCTGGAATAGAAATATTC 2565
Db      2259  CCAGACCGGAGGCCAACCTTGCAGAGCTGGAGCTACAGCAACTGGAGATAGAACTGTTT 2318
Qy      2566  CTCACAGAGCTGTCCAGACATTTTCTGTGAGGAGCAGGACAAATCTCTCTGTCCTC 2625
Db      2319  CTCACAGCTGTCTCCAGCCGTTCTCTGTGAGGACAGGAACAAATCTCTCTGTCCTC 2378
Qy      2626  AGCATCGACAGCTCTCCCTGTGGATGACAGTGAGAAACCCATCAAAGGCCCTCTCAA 2685
Db      2379  AGCATCGACAGCTCTCGCTGTGATGACAGTGAGAAAGAAATCCATCGAAGGCCCTCTCAG 2438
Qy      2686  AGCTCGGGGACACCTGAGCTCAGGCTGACGCTGACACTCTGAGCAGCAGCCGAGGCTCGA 2745
Db      2439  AGCTCAGGGGACACCTGAGTTCTGGCGTGCATCTTGGAAACAGCAAGCTGAGGCAAGA 2498
Qy      2746  AGCTCCAGCTGGAACATGCTGTGCGCCCGGGGGCGCCACCGACACACCCCAAGCTATTTC 2805
Db      2499  ACCTGAGCTGACAGCAGCGCTGCGCCGGGGCGGCTTACTGACATCCCGAGCTACTTC 2558
Qy      2806  AATGTTGTAAGTCCAAATACAGTCTTTAATGTGAACACCTGCAATCCGGAGTTC 2865
Db      2559  AACGGGGTCAAGGTCCAGATCCAGTCTCTCAATGGCGAAACCTGCATATCCGGGAATTC 2618
Qy      2866  CACCGGTTCAAAGTGGAGACATCGCCTGGCATCAGCAGCCAGATCCAGCTCAGCC 2925
Db      2619  CACCGGTCAAGTGGGAGACATGCCACCGCATCAGCAGCCAGATCCAGCCCAAGCT 2678
Qy      2926  TTCAGCTTGGTCAACAAAGACGGGAGCTGTTCGTACGATCAGAGTGCAGACTCG 2985
Db      2679  TTCAGCTTGGTCAACAAAGATGACAGCTGTTCGTATGATGATGAGGTGCCAGACTCG 2738
Qy      2986  GGCATCGACCTGAGTGACACTGGCCCTGTATGGAGCTTCGCTTGGAGCTGGAGGTC 3045
Db      2739  GGCATCGACCTGAGTGACACCTGGCCCTGTATGGCAGCTTTGCTTGGACCTGGAGGTC 2798
Qy      3046  AAGCATGGCAGCTGGAGAACAGGCCCTA 3074
Db      2799  AAGCATGGTCTGAGTAACCGACCCCTA 2827
```

RESULT 8  
US-10-087-192-886  
; Sequence 886, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:

```
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 886  
; LENGTH: 73967  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-886  
  
Query Match      33.9%; Score 1556.2; DB 13; Length 73967;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 1650; Conservative 0; Mismatches 23; Indels 10; Gaps 7;  
  
Qy      2909  AGATCCCAAGCTGCAGCCTTTCAGCTTGGTCAACCAAGACGGGAGCCTCTTCGCTACGACA 2968  
Db      62243  AGATCCCAAGCTGCAGCCTTTCAGCTTGGTCAACCAAGACGGGAGCCTCTTCGCTACGACA 62302  
  
Qy      2969  TGGAGTGCCAGACTCGGGCATTCGACCTGCGAGTGACACTGCGCCCTCATGGCAGCTTCG 3028  
Db      62303  TGGAGTGCCAGACTCGGGCATTCGACCTGCGAGTGACACTGCGCCCTCATGGCAGCTTCG 62362  
  
Qy      3029  CCTGGAGCTGAGGGGTCAAGCATGCGCAGCTGGAGAACAGCCCTTAACCCCTGCCCTCCAC 3088  
Db      62363  CCTGGAGCTGAGGGGTCAAGCATGCGCAGCTGGAGAACAGCCCTTAACCCCTGCCCTCCAC 62422  
  
Qy      3089  CGCGGCTCCACACTGCGGAAAGCAGCCTTTCCTGCTCGTGCGACGATGCTGCCCTGAAA 3148  
Db      62423  CGCGGCTCCACACTGCGCGG--AAGCAGCCTTTCCTGCTCGGTGCAGATGCTGCCCTG-AA 62480  
  
Qy      3149  ACACAGGCTCAGCGTTCCTCAGGGGATTCGACAGCCCGCCCGGCTCAGTGGGAACAGGG 3208  
Db      62481  ACACAGGCTCAGCGTTCCTCAGGGGATTCGACAGCCCGGCTCAGTGGGAACAGGG 62539  
  
Qy      3209  CCTGCAGCAGCAGAGTGGGGGCAAGCAGAGATGCTCCAGAGATTTACACCTGAGCCCT 3268  
Db      62540  CCTGCAGCAGCAGAGTGGGGGCAAGCAGAGATGCTCCAGAGATTTACACCTGAGCCCT 62599  
  
Qy      3269  GCGCCACCTGCTGAAAAACA--TCCGCCACGTGAAGAGACAGAAAGGAGGATGGCAGGAG 3327  
Db      62600  GCGCCACCTGCTGAAAAACA--TCCGCCACGTGAAGAGACAG--AGGAGGATGGCAGGAG 62658  
  
Qy      3328  TTACTCTGGGAAAAACAACAGGGATCTTTTCTGCGCCCTGCTCCAGTGCAGTTGGCCCTGA 3387  
Db      62659  TTACTCTGGGAAAAACAACAGG---ATCTTCTCTGCGCCCTGCTCCAGTGCAGTTGGCCCTGA 62714  
  
Qy      3388  CCGCTTGGATCAGTGACCATTTGTGGCAGACAGGGGAGAGCAGCTTCCAGCCTGGGTC 3447  
Db      62715  CCGCTTGGATCAGTGACCATTTGTGGCAGACAGGGGAGAGCAGCTTCCAGCCTGGGTC 62774  
  
Qy      3448  AGAAGGGGTGGGCGAGCCCTTCGCGCCCTCAACCTCCAGGCTGTGTGAGAGTGTCAAAGT 3507  
Db      62775  AGAAGGGGTGGGCGAGCCCTTCGCGCCCTCAACCTCCAGGCTGTGTGAGAGTGTCAAAGT 62834  
  
Qy      3508  GTGTAAAGGCCCAAACTCAGGTTTCAAGTGACAAACAGGTCAGAGGATGCGCCCGCTA 3567  
Db      62835  GTGTAAAGGCCCAAACTCAGGTTTCAAGTGACAAACAGGTCAGAGGATGCGCCCGCTA 62894  
  
Qy      3568  GGTAAAGGGGGCCCTCTAAACCCCTTGCTGGCCTCACCTGCCAGCTCAACCCCTTTTGG 3627  
Db      62895  GGTAAAGGGGGCCCTCTAAACCCCTTGCTGGCCTCACCTGCCAGCTCAACCCCTTTTGG 62954  
  
Qy      3628  GTGTAGGGAAAAAAGATGCTGACCCCTGGGAAGGCTTCCCTGGTAGAATACACCACTTT 3687  
Db      3628  GTGTAGGGAAAAAAGATGCTGACCCCTGGGAAGGCTTCCCTGGTAGAATACACCACTTT
```





Db 781 TATCAGTGTCCAGCGTGTGGTTCCCGAGAGCACAGCTCAGCATCACACTGACTCAC 840  
QY 4055 CTGCCCTGCCCTGGCCAGAGGTACTTGCAGCGGCACCTTTTGACTCTGATGACCTCAA 4114  
Db 841 CTGCCCTGCCCTGGCCAGAGGTACTTGCAGCGGCACCTTTTGACTCTGATGACCTCAA 900  
QY 4115 GCATTTTCATGGCTGCCCTCTGGCAGGGCAGGGCAGGTGACACACTGTAGGAGCATAG 4174  
Db 901 GCATTTTCATGGCTGCCCTCTGGCAGGGCAGGGCAGGTGACACACTGTAGGAGCATAG 960  
QY 4175 CAAGCCAGAGATGGGGTGAAGGACACAGTCTTGAGCTGTGTCACATGATGTACTCT 4234  
Db 961 CAAGCCAGAGATGGGGTGAAGGACACAGTCTTGAGCTGTGTCACATGATGTACTCT 1020  
QY 4235 CAAACCTCTTCCAGATTCTCTAAGAAATAGACCCCTTCCCATTTGCCCGAGCTAGCC 4294  
Db 1021 CAAACCTCTTCCAGATTCTCTAAGAAATAGACCCCTTCCCATTTGCCCGAGCTAGCC 1080  
QY 4295 TCTTCTCCAGGGGAGCTACTCAGGACTCACGTAGCATTAATTAATCAGCTGTGAATCGT 4354  
Db 1081 TCTTCTCCAGGGGAGCTACTCAGGACTCACGTAGCATTAATTAATCAGCTGTGAATCGT 1140  
QY 4355 GGGGTGTCTGACTAGCTCAACCTCTGGGGCAGGGGAGCCCGAGACTCCGTGGGAGAAC 4414  
Db 1141 GGGGTGTCTGACTAGCTCAACCTCTGGGGCAGGGGAGCCCGAGACTCCGTGGGAGAAC 1200  
QY 4415 TCATTCCACATCTTGCCAGAGACGCTTTGTGCCAGCTGTCCACATTGAGTCAGACTGCT 4474  
Db 1201 TCATTCCACATCTTGCCAGAGACGCTTTGTGCCAGCTGTCCACATTGAGTCAGACTGCT 1260  
QY 4475 CCCGGGAGAGACCCCGGCCCGCCAGCACATAAGAACTGACGCTTGGTACTGCGAGAGT 4534  
Db 1261 CCCGGGAGAGACCCCGGCCCGCCAGCACATAAGAACTGACGCTTGGTACTGCGAGAGT 1320  
QY 4535 CTGGGTGTAGAGAACTTTTGTAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAAA 4594  
Db 1321 CTGGGTGTAGAGAACTTTTGTAGCAATAAAGTTTGGGGTGATGACAAATGTTAAAAA 1380  
QY 4595 AA 4596  
Db 1381 AA 1382

RESULT 10  
US-10-087-192-883  
; Sequence 883, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 883  
; LENGTH: 42566  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(42566)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-883

Query Match 10.4%; Score 480.2; DB 13; Length 42566;  
Best Local Similarity 74.3%; Pred. No. 1.6e-132;

Matches 687; Conservative 0; Mismatches 153; Indels 85; Gaps 3;  
QY 554 ACAGCCAGCTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCTACAGAGGCA 613  
Db 13035 ACAGCCAGCTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCTACAGAGGCA 13094  
QY 614 AAATGGCCCGTGTGTGTGGAGGGAAGCGTCGCAGCAAAAGCCGGAAGAAACCGAAGA 673  
Db 13095 AAATGGCCCGTGTGTGTGGAGGGAAGCGTCGCAGCAAAAGCCGGAAGAAACCGTAGGA 13154  
QY 674 AGAAGAGCTCAAAAGTCCCTGGCTCATGCAGAGTGGCTTGGCCAAAACCCCTCCCCAGGA 733  
Db 13155 AGAAGAGTCAAAAGTCACTGGCCAGGAGGAGTGGCTTAGCCAAAGCCCTGCCAGAA 13214  
QY 734 CCCCTGAGCAGGAGAGCTGCACCATCCAG----- 763  
Db 13215 CCCCTGAGCAAGAGAGCTGTACCATCCCAAGTACAGGTGAGATCCTTGCCCGCTGGT 13274  
QY 764 ----- 774  
Db 13275 CCCAGCTCATGGCGTTTGGAGCTGGGACTTGTAAATGCCTGTATTTTCTCGCAGGAAGAT 13334  
QY 775 GAGTCTCACTGGCGCCCCATATGTTAGAAAACACCCCGCAGTTTACCAAGCCTCTGAAG 834  
Db 13335 GAGTCTCACTAGGCAACCTCTATGCCAGAAAATGTCTCCAGTTTCAACAAGCCTCTGGG 13394  
QY 835 GAAACAGGCTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCTACGGCCGCTCTG 894  
Db 13395 GAAACAGGCTTGGGCAACTCTGTTTTAAGAAAACAAAGATGAAGGCTTCGACCCGCTACTG 13454  
QY 895 CCTCGATCAGAACTCCACAAACTGATCAGCCCTTGCATATGTCTGAAACACAGTGTGAAA 954  
Db 13455 CCTCGACAGAACTCCACAAACTGATCAGCCCTTGCATATGTCTTAAACACAGTGTGAAA 13514  
QY 955 CTGACACACCCCGAGGAGGCCCCCTGCCCTTGCACCCAGCCACCCCTTCCCCTATAGC 1014  
Db 13515 CTCCACACACCCCGAGGACACAGGCCCCCGGCCACCCGACTCACCCCTTCCCCTACAGC 13574  
QY 1015 AGACTGCCCTCATCCCTTCCGATTCACCCCTTCCAGCCCTCGAAACCTTCAACCTCTGGAG 1074  
Db 13575 GGAATGCCCCCATCCTTTTCCCATTTACCCCTTGGAGCCCTGGAACCCCTATATGCTGGAC 13634  
QY T---CCTTCTCGGCAAACTGGCCTGTGTAGACAGCCAGAAAACCTTGCCTGACCCA--- 1128  
Db 13635 TCTGCCGTCTCGACAAACTAGCCGGTGTACGCGGCGCAGCGCCCTCTGCTTGGCCCAACG 13694  
QY 1129 CACTGAGCAAACTGGCCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCAACCTGGAG 1188  
Db 13695 CATCTAAGCCAACTGGCCCATGGAGACAGTCAAGAGCCGCTGCCTGGCCCAACCTGGAG 13754  
QY 1189 CCCAGCTGCCCTGTCTCGTGGTGCCCATGAGAAAGTTTCTGTGGAGGAATACCTAGTGAT 1248  
Db 13755 TCCAGCTGCCCTGTCTCGGGGTGCCCTAGAAAAGGTTCCTGGAGGAATACCTAGTGAT 13814  
QY 1249 GCTCTGCAAGCAGCGTGTAGCTCAAGCCAGGCCCCACACAGCTTGCAGCCTTGCCCAAGACC 1308  
Db 13815 GCGCTCAAGAAAGTGTGAGCTCAGGCCAGGCCCCACACAGCTTGGCAGCTTGGCTAAGACA 13874  
QY 1309 TGGGAGCAGCGGGCTCCAGATCCCGGAGCCCGAGCCCAAACTGAGGACAAAGAGGT 1368  
Db 13875 TGGTCTCGGGAAGCGCCAGCTGCAGAGGCTCGGCCCGGAAACTGAGGACAAACGAGGGG 13934  
QY 1369 GTCCCTGCTCACTGAGAACTCAAGC 1393  
Db 13935 GTCCCTGCTTACTGAGTAACCTAACC 13959

RESULT 11  
US-10-108-260A-1754/c  
; Sequence 1754, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE



```
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1754
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1754

Query Match          7.6%; Score 349.4; DB 17; Length 2475;
Best Local Similarity 99.7%; Pred. No. 1.1e-93;
Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2200 CAAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCAAGACATCCACCG 2259
Db 2179 CTAGTGGAGGTCTGAAGAGCCCTTGGAGGGGAGAAATAAAGAACCAAGACATCCACCG 2120
Qy 2260 CCAATCAAGCAATTACCAAGACCTTCCATGCCAGCCGAGAGACTTTCGCCAAGG 2319
Db 2119 CCAATCAAGCAATTACCAAGACCTTCCATGCCAGCCGAGAGACTTTCGCCAAGG 2060
Qy 2320 GCCCCAGGGCCCCGCGAGCTGAGGAGACAACAGCAGAGCCCTTAAGCTCCAGCTCCT 2379
Db 2059 GCCCCAGGGCCCCGCGAGCTGAGGAGACAACAGCAGAGCCCTTAAGCTCCAGCTCCT 2000
Qy 2380 CTCACCACAGAGCCCCGAGAGCAAAACAAGTCTCTCTCTTGTGAGCAAGAGGAG 2439
Db 1999 CTCACCACAGAGCCCCGAGAGCAAAACAAGTCTCTCTCTTGTGAGCAAGAGGAG 1940
Qy 2440 TCTGGATGTGGGAACCCCTTACCTCTGTCCTCTCTGGAGCCAGCCCTTCCAGAAACCCC 2499
Db 1939 TCTGGATGTGGGAACCCCTTACCTCTGTCCTCTCTGGAGCCAGCCCTTCCAGAAACCCC 1880
Qy 2500 AGCTCACCAGAGCGGAAGCAACCGTCCGAGCAGAGCACTGCAGCAGCTG 2550
Db 1879 AGCTCACCAGAGCGGAAGCAACCGTCCGAGCAGAGCACTGCAGCAGCTG 1829

RESULT 12
US-10-433-793-141
; Sequence 141, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 141
; LENGTH: 6673
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-141

Query Match          6.3%; Score 291.6; DB 19; Length 6673;
Best Local Similarity 76.3%; Pred. No. 3.7e-76;
Matches 371; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

Qy 5 GGGGAGCTGCTCCCTGTGGACGTGTAGCTTGA-AGGTGACTCTGTACCATTCAGG 63
Db 5005 GGGGAGTGTCTGTGTGGAACGTGTAGTGTGTGAGAGGTGGATTTGTATTATTAGG 5064
Qy 64 ATGTTTGGAGGATGAGTATGTGTGGCAGAGCAGACACATATAACAGCAGACACCTTTGCC 123
Db 5065 ATGTTTGGAGGATGAGTATGTGTGGTAGAGGTATATATAATAGTAGAGATTTTGT 5124
Qy 124 CCTGCCTTTCTCCCCCAACCAAGGCTGACCTGTGTCTTCTCCAGGTCTGGGATTTCTAAGT 183
```

---

```
Db 5125 TTGTGTTTTTTTTTTTAAATTAAGTTGATTTGTTTTTTTAGGATTTTAAAGT 5184
Qy 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGACACAAGCCTGGGAGATGGCAGTGTATG 243
Db 5185 GATTTGTTTGTGTTTGGTCTCTCTCAGGATGACACAAGTGTGGAGATGGTGTATGATG 5244
Qy 244 GAAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCAAG 303
Db 5245 GAAATGGTGTGTTTGGTCTCTCTCAGGATGACACAAGTGTGGAGTGTGGAGTGTG 5304
Qy 304 CCAAGGAGAGAGACGCCGACCTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 363
Db 5305 GTTAAGGAGAGAGAGCTGCTTATTGTTGGGAAGAAATAGAGTTTCGTTTATAAGTTTGAGGTC 5364
Qy 364 GTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCCTGAATGAGCTGATTTACCAAG 423
Db 5365 GTGAGAGAGAGTGTGTTGTTTTCGGAAGTGGGAGATTTTGAATGAGCTGATTTATTAAG 5424
Qy 424 GGCAAGCAAGAGAGGCTCCGAGGAGGCGCAGCTGCTATCTATCATCGCCAGGCT 483
Db 5425 GGTATAGTTAAGGAAGGTTTCGAGGTAGGTTAGTTGTTTATTTATTCGTTTAGGTT 5484
Qy 484 GAGTGT 489
Db 5485 GAGTGT 5490

RESULT 13
US-09-918-995-2585
; Sequence 2585, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2585
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(499)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-2585

Query Match          5.8%; Score 267.4; DB 10; Length 499;
Best Local Similarity 95.6%; Pred. No. 2.8e-69;
Matches 307; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

Qy 2909 AGATCCAGCTGCAGCCTTCAGCTTGTTCACCAAGAGCGGCGAGCCTGTTTCGTACGACA 2968
Db 182 AGATCCAGCTGCAGCCTTCAGCTTGTTCACCAAGAGCGGCGAGCCTGTTTCGTACGACA 241
Qy 2969 TGGAGGTGCCAGACTCGGGGATCGACCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTCG 3028
Db 242 TGGAGGTGCCAGACTCGGGGATCGACCTGCAGTGCACACTGGCCCCCTGATGGCAGCTTCG 301
Qy 3029 CTTGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCAC 3088
Db 302 CTTGAGCTGGAGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTTAACCTGCCCTCCAC 361
Qy 3089 CGCCGGCTCCACACTCGCCGAAAGCAGCCTTCTGCTCGGTGACGATGCTGCCCTGAAA 3148
Db 362 CGCCGGCTCCACACTCGCCG-AMGACGCTTCTGCTCGGTGACGATGCTGCCCTG-AA 419
```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2005, 23:21:54 ; Search time 506.684 Seconds  
(without alignments)  
14842.245 Million cell updates/sec

Title: US-09-155-676B-6

Perfect score: 4596

Sequence: 1 agcgggggactgtgcgtg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 8181359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3149.6	68.5	3156	2	US-08-887-518-1
2	3149.6	68.5	3156	2	US-09-023-321-1
3	3149.6	68.5	3156	2	US-09-032-475-1
4	2837.6	61.7	2844	3	US-09-257-703-2
5	2837.6	61.7	2844	4	US-09-871-889A-2
6	288.2	6.5	362	4	US-09-513-999C-2583
7	88.8	1.9	2207	4	US-09-949-016-5071
8	88.8	1.9	2348	4	US-09-658-688A-3
9	83.8	1.8	3089	1	US-08-472-934-5
10	83.8	1.8	3089	2	US-08-323-460A-5
11	83.8	1.8	3089	2	US-08-461-146C-5
12	83.8	1.8	3089	3	US-08-461-145C-5
13	83.8	1.8	3089	3	US-08-628-829-9
14	83.8	1.8	3332	3	US-09-423-890-11
15	82.2	1.8	1935	3	US-09-423-890-5
16	61.6	1.3	3260	1	US-08-049-254-1
17	61.6	1.3	3260	1	US-08-472-934-1
18	61.6	1.3	3260	2	US-08-323-460A-1
19	61.6	1.3	3260	2	US-08-461-146C-1
20	61.6	1.3	3260	3	US-08-461-145C-1
21	61.6	1.3	3260	3	US-08-628-829-1
22	61.6	1.3	5253	3	US-09-423-890-7
23	61.6	1.3	5539	3	US-08-628-829-3
24	60.8	1.3	8563	4	US-09-902-540-3318
25	60.8	1.3	15351	4	US-09-902-540-1154
26	60.4	1.3	9862	4	US-09-691-861A-3
27	59.4	1.3	601	4	US-09-691-861A-14

28	58.8	1.3	1104	4	US-09-691-861A-1	Sequence 1, Appli
29	58.6	1.3	2169	4	US-09-016-434-1147	Sequence 1147, Ap
30	58.6	1.3	4162	2	US-08-459-448A-26	Sequence 26, Appl
31	58.6	1.3	4162	3	US-08-459-595A-26	Sequence 26, Appl
32	58.6	1.3	4162	3	US-08-459-504B-26	Sequence 26, Appl
33	58.6	1.3	4162	3	US-08-459-444-26	Sequence 26, Appl
34	58.6	1.3	4162	3	US-09-547-422-26	Sequence 26, Appl
35	58.6	1.3	4162	4	US-09-988-462-26	Sequence 26, Appl
36	58.6	1.3	4165	1	US-07-951-715A-26	Sequence 26, Appl
37	57.8	1.3	1732	3	US-09-430-564-1	Sequence 1, Appli
38	57.8	1.3	1816	4	US-09-762-258-1	Sequence 1, Appli
39	57.2	1.2	1458	4	US-09-230-896C-5	Sequence 5, Appli
40	57	1.2	320	4	US-09-270-767-536	Sequence 536, App
41	57	1.2	320	4	US-09-270-767-15818	Sequence 15818, A
42	56.4	1.2	1442	4	US-09-949-016-1590	Sequence 1590, Ap
43	56.4	1.2	1442	4	US-09-949-016-1591	Sequence 1591, Ap
44	56.4	1.2	1480	4	US-09-016-434-1454	Sequence 1454, Ap
45	55.2	1.2	1282	2	US-08-878-989-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1  
US-08-887-518-1  
; Sequence 1, Application US/08887518  
; Patent No. 5843721  
; GENERAL INFORMATION:  
; APPLICANT: Rothe, Mike  
; APPLICANT: Wu, Lin  
; TITLE OF INVENTION: NIK Proteins, Nucleic Acids and Methods  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/887,518  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: T97-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3156 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; US-08-887-518-1

Query Match 68.5%; Score 3149.6; DB 2; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	232	ATGCGAGTGTGGAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGCGACAGAG 291
Db	1	ATGCGAGTGTGGAATGGCTGCCAGGTGCCCTCGCTCAGAGTGGGCGACAGAG 60
Qy	292	GAACCTCCCAAGCCAAAGAGAGACGCCCGCCACTGGGGAGAAACAGAGCTCGCTCTAC 351

Db 61 GAACTCCCAAGACCAAGAGAGAGCGCGCCACTGGGAGAAACAGAGCTCGTCTAC 120  
QY 352 AAGCTTGAGGCGGTGAGAGAGAGCCCTGTGTTCGCGAAAGTGGAGATCCTGAATGAC 411  
Db 121 AAGCTTGAGGCGGTGAGAGAGAGCCCTGTGTTCGCGAAAGTGGAGATCCTGAATGAC 180  
QY 412 GTGATTACAGGGGACAGCCAGAGAGGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 471  
Db 181 GTGATTACCAAGGGACAGCCAAAGAGGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 240  
QY 472 ATCGGCCAGGCTGAGTGTGAGAAATAGCCAAGAGTTACGCCCCACCTTTTCAGAAAGCAATT 531  
Db 241 ATCGGCCAGGCTGAGTGTGAGAAATAGCCAAGTTCAGCCCCACCTTTTCAGAAAGCAATT 300  
QY 532 TTCAATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTGTGATCAGATCCCAACAAT 591  
Db 301 TTCAATCGCTGGTCCAAACAGTACAGCCAGTCCGAGAGTCTGTGATCAGATCCCAACAAT 360  
QY 592 GTGGCCCATGCTACAGAGGGCAAAATGCCCGTGTGTGTTGGAAGGAAAGCGTCGCAGC 651  
Db 361 GTGGCCCATGCTACAGAGGGCAAAATGCCCGTGTGTGTTGGAAGGAAAGCGTCGCAGC 420  
QY 652 AAAGCCCGAAGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCGAGGAGTGCC 711  
Db 421 AAAGCCCGAAGAAACGGAAGAGAGCTCAAAAGTCCCTGGCTCATGCGAGGAGTGCC 480  
QY 712 TTGGCCAAACCCCTCCCAAGACCCCTGAGCGAGGAGGTGACCAATCCCAAGTGAGGAG 771  
Db 481 TTGGCCAAACCCCTCCCAAGACCCCTGAGCGAGGAGGTGACCAATCCCAAGTGAGGAG 540  
QY 772 GATGAGTCTCACTGGCGGCCCATATGTTAGAAACACCCGAGTTCACCAAGCCTCTG 831  
Db 541 GATGAGTCTCACTGGCGGCCCATATGTTAGAAACACCCGAGTTCACCAAGCCTCTG 600  
QY 832 AAGGAACCAAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCT 891  
Db 601 AAGGAACCAAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCGGCT 660  
QY 892 CTGCTCTGATCAGAACTCCAAACCTGATCAGCCCTTTGCAATGTCTGAACCACTGTGG 951  
Db 661 CTGCTCTGATCAGAACTCCAAACCTGATCAGCCCTTTGCAATGTCTGAACCACTGTGG 720  
QY 952 AAACCTGACACCCCGAGAGCGGCGCCCTGCCCTGCCACGACCCCTTCCCTTA 1011  
Db 721 AAACCTGACACCCCGAGAGCGGCGCCCTGCCCTGCCACGACCCCTTCCCTTA 780  
QY 1012 AGCAGACTGCCTCATCCCTTCCCATTTCCACCCCTCTCCAGCCCTTGAAACCTCACCCCTCTG 1071  
Db 781 AGCAGACTGCCTCATCCCTTCCCATTTCCACCCCTCTCCAGCCCTTGAAACCTCACCCCTCTG 840  
QY 1072 GAGTCTTCTGGGCAAACTGGGCTGTGTAGACAGCGAGAAACCTTGGCTGACCCACAC 1131  
Db 841 GAGTCTTCTGGGCAAACTGGGCTGTGTAGACAGCGAGAAACCTTGGCTGACCCACAC 900  
QY 1132 CTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCAACCTTGAGGCC 1191  
Db 901 CTGAGCAAACTGCGCTGTGTAGACAGTCCAAAGCCCTTGCCTGGCCCAACCTTGAGGCC 960  
QY 1192 AGCTGCTGTCTGTGTGCCATGAGAAATTTCTGTGGAGGAAATACCTAGTGCATGCT 1251  
Db 961 AGCTGCTGTCTGTGTGCCATGAGAAATTTCTGTGGAGGAAATACCTAGTGCATGCT 1020  
QY 1252 CTGCAAGGAGGCTGAGCTCAAGCGAGGCCACAGCTGACAGCCCTGGCCAGAGCCTGG 1311  
Db 1021 CTGCAAGGAGGCTGAGCTCAAGCGAGGCCACAGCTTGACCCCTGGCCAGAGCCTGG 1080  
QY 1312 GCAGCAGCGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGTGTC 1371  
Db 1081 GCAGCAGCGGCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAACGAGGTGTC 1140  
QY 1372 CTGCTCACTGAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCC 1431

Db 1141 CTGCTCACTGAGAAACTCAAAGCCAGTGGATTATGATGATACCGAGAAGATCCACTGGGCC 1200  
QY 1432 ACGCACAGCTCCGCTGGCGAGAGGCTCCTTCGAGAGGTGCACAGATGGAGGACAAG 1491  
Db 1201 ACGCACAGCTCCGCTGGCGAGAGGCTCCTTCGAGAGGTGCACAGATGGAGGACAAG 1260  
QY 1492 CAGACTGGCTTCAGTGCCTGTCAAAGAGTGCCTCGCTGCGAAAGTATTTCCGGCGCAGAGGAG 1551  
Db 1261 CAGACTGGCTTCAGTGCCTGTCAAAGAGTGCCTCGCTGCGAAAGTATTTCCGGCGCAGAGGAG 1320  
QY 1552 CTGATGGCATGTGAGGATTTGACCTCAACCAGAAATTTCTCCCTTTGTATGAGAGTGTGAGA 1611  
Db 1321 CTGATGGCATGTGAGGATTTGACCTCAACCAGAAATTTCTCCCTTTGTATGAGAGTGTGAGA 1380  
QY 1612 GAAGGGCTTTGGTCAACATCTTCATGAGAGTGCCTGGAAGTGGCTCCCTTGGGCCAGCTG 1671  
Db 1381 GAAGGGCTTTGGTCAACATCTTCATGAGAGTGCCTGGAAGTGGCTCCCTTGGGCCAGCTG 1440  
QY 1672 GTCAAGGAGCAGGCGCTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1731  
Db 1441 GTCAAGGAGCAGGCGCTCTCCAGAGGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTG 1500  
QY 1732 GAGGCTGTGAATTAACCTCCACTCAAGAAAGATTTCTGATGGGACGTCAAAGCTGACAAC 1791  
Db 1501 GAGGCTGTGAATTAACCTCCACTCAAGAAAGATTTCTGATGGGACGTCAAAGCTGACAAC 1560  
QY 1792 GTGCTCTGTCCAGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGTCTGTGT 1851  
Db 1561 GTGCTCTGTCCAGCGATGGAGCCACGACGCCCTCTGTGACTTTGGCCATGTCTGTGTGT 1620  
QY 1852 CTTCAACCTGATGCCCTGGGAAAGTCTCTGTCTCAAGGGGACTACATCCCTGGGCACAGAG 1911  
Db 1621 CTTCAACCTGATGCCCTGGGAAAGTCTCTGTCTCAAGGGGACTACATCCCTGGGCACAGAG 1680  
QY 1912 ACCCAATGGCTCCGAGGTGTGCTGGCGAGGAGTGCAGCGCCAGAGTGGATGTCTGG 1971  
Db 1681 ACCCAATGGCTCCGAGGTGTGCTGGCGAGGAGTGCAGCGCCAGAGTGGATGTCTGG 1740  
QY 1972 AGCAGCTGTGTATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTCTTC 2031  
Db 1741 AGCAGCTGTGTATGATGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTCTTC 1800  
QY 2032 CGAGGGCGCTCTGCCTCAAGATTTGCCAGGAGCCTCCGCTGTGAGGGAGATCCCAACC 2091  
Db 1801 CGAGGGCGCTCTGCCTCAAGATTTGCCAGGAGCCTCCGCTGTGAGGGAGATCCCAACC 1860  
QY 2092 TCCTGGGCCCTCTCAAGCCAGGCCATCAAAGGGGCTGAGGAAAGAGCCCATCCAC 2151  
Db 1861 TCCTGGGCCCTCTCAAGCCAGGCCATCAAAGGGGCTGAGGAAAGAGCCCATCCAC 1920  
QY 2152 CGCTGTCTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGT 2211  
Db 1921 CGCTGTCTGCAGCGGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGT 1980  
QY 2212 CTGAAGAGCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2271  
Db 1981 CTGAAGAGCCCTTGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCC 2040  
QY 2272 AATTACCAACAGACCTTCATGCCAGCCGAGAGGCTTCGCCAAGGGCCCAAGGGCCC 2331  
Db 2041 AATTACCAACAGACCTTCATGCCAGCCGAGAGGCTTCGCCAAGGGCCCAAGGGCCC 2100  
QY 2332 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAG 2391  
Db 2101 CGGCCAGCTGAGGAGACAAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACACAGAG 2160  
QY 2392 CCCCCAGAGCCAAACAAAGTCTCTCTCCCTTGAATTGTGACAAAGGAGATGTGGGATGTGG 2451  
Db 2161 CCCCCAGAGCCAAACAAAGTCTCTCTCCCTTGAATTGTGACAAAGGAGATGTGGGATGTGG 2220  
QY 2452 GAAACCTTACCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2511  
Db 2221 GAAACCTTACCTGTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAG 2280

Qy	2512	CGGAAAGCAACGTTCCCGAGCAGGAAACTGCAGCAGCTGGAAATATTTCCTCAAC	2571
Db	2281	CGGAAAGCAACGTTCCCGAGCAGGAAACTGCAGCAGCTGGAAATATTTCCTCAAC	2340
Qy	2572	AGCCTGTCAGGCAATTTTCTCTGGAGGAGCAGGAGCAAAATTCCTCGTGCCTCAGCATC	2631
Db	2341	AGCCTGTCAGGCAATTTTCTCTGGAGGAGCAGGAGCAAAATTCCTCGTGCCTCAGCATC	2400
Qy	2632	GACAGCCTCTCCCTGTCCGAGTGACAGTGAGAAGAACCCATCAAAAGGCTCTCTCAAAGCTCG	2691
Db	2401	GACAGCCTCTCCCTGTCCGAGTGACAGTGAGAAGAACCCATCAAAAGGCTCTCTCAAAGCTCG	2460
Qy	2692	CGGGAACCCCTGAGCTCAGGCGTACATCTCTGGAGCAGCGAGGCCGAGGCTCGAAGTCTC	2751
Db	2461	CGGGAACCCCTGAGCTCAGGCGTACATCTCTGGAGCAGCGAGGCCGAGGCTCGAAGTCTC	2520
Qy	2752	AGCTGGAAACATGGTGTGGCCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGGT	2811
Db	2521	AGCTGGAAACATGGTGTGGCCCGGGGGGCCACCGACACCCCAAGCTATTTCAATGGT	2580
Qy	2812	GTGAAAGTCCAAATACAGTCTCTTAATGGTGAAACACTGCACATCCGGGAGTTCCACCGG	2871
Db	2581	GTGAAAGTCCAAATACAGTCTCTTAATGGTGAAACACTGCACATCCGGGAGTTCCACCGG	2640
Qy	2872	GTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAGCTCGAGCCTTCAGC	2931
Db	2641	GTCAAAGTGGGAGACATCGCCACTGGCATCAGCAGCCAGATCCCAGCTCGAGCCTTCAGC	2700
Qy	2932	TTGGTGACAAAGACGGGCGACCTGTTTCGGTACGACATGGAGGTGCAGACTCGGGCATC	2991
Db	2701	TTGGTGACAAAGACGGGCGACCTGTTTCGGTACGACATGGAGGTGCAGACTCGGGCATC	2760
Qy	2992	GACCTGAGTGCACACTGGGCCCTGTATGGCGACTTTCGCTGGAGCTGGAGGGTCAAGCAT	3051
Db	2761	GACCTGAGTGCACACTGGGCCCTGTATGGCGACTTTCGCTGGAGCTGGAGGGTCAAGCAT	2820
Qy	3052	GGCCAGCTGAGAAACAGCGCCCTAACCCCTGCCTCCACCGCGGCTCCACACTGCCGGA	3111
Db	2821	GGCCAGCTGAGAAACAGCGCCCTAACCCCTGCCTCCACCGCGGCTCCACACTGCCGGA	2880
Qy	3112	GCAGCCTTCTGCTCGGTGACGATGTCGCCCTGAAACACAGGCTCAGCCGTTCCGAGG	3171
Db	2881	GCAGCCTTCTGCTCGGTGACGATGTCGCCCTGAAACACAGGCTCAGCCGTTCCGAGG	2940
Qy	3172	GGATTGCCAGCCCCCGGCTCAGTGTGGAAACAGAGGCCCTCGCAGCAGCAGGTGGGGG	3231
Db	2941	GGATTGCCAGCCCCCGGCTCAGTGTGGAAACAGAGGCCCTCGCAGCAGCAGGTGGGGG	3000
Qy	3232	AAGCAGAAATGCCCTCCAGGATTTTCACACCTGAGCCCTGCCCACTGCTGAAAAACAT	3291
Db	3001	AAGCAGAAATGCCCTCCAGGATTTTCACACCTGAGCCCTGCCCACTGCTGAAAAACAT	3060
Qy	3292	CCGCCAGTGAAGACAGAAAGGAGTGGCAGGAGTTACTGTGGGGAACAAACACGGGA	3351
Db	3061	CCGCCAGTGAAGACAGAAAGGAGTGGCAGGAGTTACTGTGGGGAACAAACACGGGA	3120
Qy	3352	TCTTTTTCTGCCCTGTCTCAGTCGAGTTGGCCTGA	3387
Db	3121	TCTTTTTCTGCCCTGTCTCAGTCGAGTTGGCCTGA	3156

## RESULT 2

RESUL 2  
US-09-023-321-1

US-09-023-321-1  
: Sequence 1, Application US/09023321

; Sequence 1, Application No. 5844073

GENERAL INFORMATION:

APPLICANT: Rothe, Mike

APPLICANT: Wu, Lin

**TITLE OF INVENTION:** NIK Proteins, Nucleic Acids and Methods

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,321  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/887,518  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A.  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: T97-008  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3156 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-09-023-321-1

Query Match 68.5%; Score 3149.6; DB 2; Length 3156;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

232	ATGGCAGTGTGAAATGGCTGCCCAAGTGTGCCCTGGCTCAGCGATGGGGCAGCAGAG	291
1	ATGGCAGTGTGAAATGGCTGCCCAAGTGTGCCCTGGCTCAGCAGTGGGGCAGCAGAG	60
292	GAACTCCCCAAGCACAAGGAGAGACCGCCCACTGGGGAAGAAACAGAGTCCGCTTAC	351
61	GAACTCCCCAAGCACAAGGAGAGACCGCCCACTGGGGAAGAAACAGAGTCCGCTTAC	120
352	AAGCTTGAGGCCGTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGAC	411
121	AAGCTTGAGGCCGTGGAGAAAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGAC	180
412	GTGATTACCAAGGGCACAGCCAAGGAAGGCTCCGAGGCAGGGCCAGCTGCCATCTCTATC	471
181	GTGATTACCAAGGGCACAGCCAAGGAAGGCTCCGAGGCAGGGCCAGCTGCCATCTCTATC	240
472	ATGCCCCAGGCTGAGTGTGAAATAGCAAGAGTTCAGCCCCACCTTTTCAGAAACGCATT	531
241	ATGCCCCAGGCTGAGTGTGAAATAGCAAGAGTTCAGCCCCACCTTTTCAGAAACGCATT	300
532	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAT	591
301	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAT	360
592	GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTAAGGGAAAGCGTCGCAGC	651
361	GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTAAGGGAAAGCGTCGCAGC	420
652	AAAGCCCGGAAGAAAACGGAAGAAGAGAGCTCAAAGTCCCTGGCTCATGSCAGGAGTGGCC	711
421	AAAGCCCGGAAGAAAACGGAAGAAGAGAGCTCAAAGTCCCTGGCTCATGSCAGGAGTGGCC	480
712	TTGGCCAAACCCCTCCCGAAGACCCCTTGAGCAGAGAGCTGCACCATCCCATGTCAGGAG	771
481	TTGGCCAAACCCCTCCCGAAGACCCCTTGAGCAGAGAGCTGCACCATCCCATGTCAGGAG	540





Db 2701 TTGCTACCAAGACGGGCGCCCTGTTGGCTACGACATGAGGTGCCAGACTCGGGCATC 2760  
Qy 2992 GACCTGCAGTGACACACTGGCCCTCGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 3051  
Db 2761 GACCTGCAGTGACACACTGGCCCTCGATGGCAGCTTCGCTGGAGCTGGAGGGTCAAGCAT 2820  
Qy 3052 GGCAGCTGGAGAACAGCCCTTAACCTTGCCTCCAGCGCGGGTCCACACTGCGGAAA 3111  
Db 2821 GGCAGCTGGAGAACAGCCCTTAACCTTGCCTCCAGCGCGGGTCCACACTGCGGAAA 2880  
Qy 3112 GCAGCCTTCCTGCTCGGTGCACAGATGCTGCCCTGAAACACAGGCTCAGCGGTTCCCGAG 3171  
Db 2881 GCAGCCTTCCTGCTCGGTGCACAGATGCTGCCCTGAAACACAGGCTCAGCGGTTCCCGAG 2940  
Qy 3172 GGATTGCGAGCCCCCGCTCACAGTGGGAACACAGGCGCTCGCAGCAGCAAGGTGGGGGC 3231  
Db 2941 GGATTGCGAGCCCCCGCTCACAGTGGGAACACAGGCGCTCGCAGCAGCAAGGTGGGGGC 3000  
Qy 3232 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCGCCACCTGCTGAAAAACAT 3291  
Db 3001 AAGCAGATGCTCCAGGATTTACACCTGAGCCCTGCGCCACCTGCTGAAAAACAT 3060  
Qy 3292 CCGCCAGTGAAGACAGAGAGGAGGATGGCAGGATTTACCTGGGAAACAAAAACAGGGA 3351  
Db 3061 CCGCCAGTGAAGACAGAGAGGAGGATGGCAGGATTTACCTGGGAAACAAAAACAGGGA 3120  
Qy 3352 TCTTTTCTGCCCCCTGCTCCAGTGCAGTTGGCCCTGA 3387  
Db 3121 TCTTTTCTGCCCCCTGCTCCAGTGCAGTTGGCCCTGA 3156

RESULT 3

US-09-032-475-1  
; Sequence 1, Application US/09032475  
; Patent No. 5854003  
; GENERAL INFORMATION:  
; APPLICANT: Rothe, Mike  
; APPLICANT: Wu, Lin  
; TITLE OF INVENTION: NFK Proteins, Nucleic Acids and Methods  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/032,475  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/887,518  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: T97-008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3156 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA

US-09-032-475-1  
Query Match 68.5%; Score 3149.6; DB 2; Length 3156;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3152; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 232 ATGGCAGTGTGATAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAG 291  
Db 1 ATGGCAGTGTGATAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAG 60  
Qy 292 GAATCTCCCAAGCCAAAGGAGAGAGCCGCCACTGGGGAGAGAAACAGAGCTCGTCTTAC 351  
Db 61 GAATCTCCCAAGCCAAAGGAGAGAGCCGCCACTGGGGAGAGAAACAGAGCTCGTCTTAC 120  
Qy 352 AAGCTTGAGGCCGTGGAGAGAGCCCTGTGTCTTCTGCGGAAAGTGGGAGATCTCTGAATGAC 411  
Db 121 AAGCTTGAGGCCGTGGAGAGAGCCCTGTGTCTTCTGCGGAAAGTGGGAGATCTCTGAATGAC 180  
Qy 412 GTGATTACCAAGGGCAGCAGCAAGGAAGGCTCCGAGGCAGGGCCAGCTGCCATCTCTATC 471  
Db 181 GTGATTACCAAGGGCAGCAGCAAGGAAGGCTCCGAGGCAGGGCCAGCTGCCATCTCTATC 240  
Qy 472 ATCGCCAGGCTCAGTGTGAGAAATAGCCAAGAGTTTTCAGCCCCACCTTTTCAGAACGCAAT 531  
Db 241 ATCGCCAGGCTCAGTGTGAGAAATAGCCAAGAGTTTTCAGCCCCACCTTTTCAGAACGCAAT 300  
Qy 532 TTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 591  
Db 301 TTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAT 360  
Qy 592 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAAAGGAAAGCGTCGGAGC 651  
Db 361 GTGGCCCATGCTACAGAGGGCAAAATGGCCCGTGTGTGTGGAAAGGAAAGCGTCGGAGC 420  
Qy 652 AAAGCCGGAAGAAACGGAAGAGAGAGCTCAAAGTCCCTGGCTCATGAGGAGTGGCC 711  
Db 421 AAAGCCGGAAGAAACGGAAGAGAGAGCTCAAAGTCCCTGGCTCATGAGGAGTGGCC 480  
Qy 712 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAG 771  
Db 481 TTGGCCAAACCCCTCCCGAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAG 540  
Qy 772 GATGAGTCTCCACTCGGGCGCCCATATGTTAGAAAACACCCCGCAGTTTCCAGCCCTCTG 831  
Db 541 GATGAGTCTCCACTCGGGCGCCCATATGTTAGAAAACACCCCGCAGTTTCCAGCCCTCTG 600  
Qy 832 AAGGAACCCAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCCGCT 891  
Db 601 AAGGAACCCAGGCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGGCCGCT 660  
Qy 892 CTGCCTCGATCAGAACTCCCAAACTGATCAGCCCTTTGCAATGCTGTGAACCAAGTGTGG 951  
Db 661 CTGCCTCGATCAGAACTCCCAAACTGATCAGCCCTTTGCAATGCTGTGAACCAAGTGTGG 720  
Qy 952 AAATGCAACACCCCGAGGAGCGGCGCCCTGCGCCCTGCGCCAGCAGCACCCCTTCCCTAT 1011  
Db 721 AAATGCAACACCCCGAGGAGCGGCGCCCTGCGCCCTGCGCCAGCAGCACCCCTTCCCTAT 780  
Qy 1012 AGCAGACTGCTCATCCCTTCCCACTTCCAGCCCTTCCAGCCCTTCCAGCCCTCTG 1071  
Db 781 AGCAGACTGCTCATCCCTTCCCACTTCCAGCCCTTCCAGCCCTTCCAGCCCTCTG 840  
Qy 1072 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTTGCCCTGACCCACAC 1131  
Db 841 GAGTCTCTCTGGGCAAACTGGGCTGTGTAGACAGCCAGCAAAACCTTTGCCCTGACCCACAC 900  
Qy 1132 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACTCTGGAGGCC 1191  
Db 901 CTGAGCAAACTGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACTCTGGAGGCC 960  
Qy 1192 AGCTGCCTGTCTGTTGGTCCCATGAGAGTCTTCTGTGGAGGAAATACCTAGTGCATGCT 1251  
Db 961 AGCTGCCTGTCTGTTGGTCCCATGAGAGTCTTCTGTGGAGGAAATACCTAGTGCATGCT 1020

QY	1252	CTGCAAGCAGCGTGAGCTCAAGCCAGGCCCCACAGCCTGACCAGCTGGGCCAAGACCTGG	1311
DB	1021	CTGCAAGCAGCGTGAGCTCAAGCCAGGCCCCACAGCCTGACCAGCTGGGCCAAGACCTGG	1080
QY	1312	GCAGCACGGGGCTCCAGATCCCGGAGGCCACGAGCCCCAAAACCTGAGGACAACGAGGGGTGC	1371
DB	1081	GCAGCACGGGGCTCCAGATCCCGGAGGCCCCAAACCTGAGGACAACGAGGGGTGC	1140
QY	1372	CTGCTCACTAGAAACTCAAGCCAGTGGATTATGATACCGAGAAGATCCACTGGGCC	1431
DB	1141	CTGCTCACTAGAAACTCAAGCCAGTGGATTATGATACCGAGAAGATCCACTGGGCC	1200
QY	1432	ACGCACAGCTCCGGCTGGGCGAGAGGCTCTTCGGAGAGGTGCACAGGATGAGAGACAAG	1491
DB	1201	ACGCACAGCTCCGGCTGGGCGAGAGGCTCTTCGGAGAGGTGCACAGGATGAGAGACAAG	1260
QY	1492	CAGACTGGCTTCCAGTGGCTGTCAAAGAGGTGGCTGGAAGTATTTCGGGCGAGAGGAG	1551
DB	1261	CAGACTGGCTTCCAGTGGCTGTCAAAGAGGTGGCTGGAAGTATTTCGGGCGAGAGGAG	1320
QY	1552	CTGATGGCAGTGCAGGATTGACTCACCCAGAAATTCTCCCTTTGTATGAGCTGTGAGA	1611
DB	1321	CTGATGGCAGTGCAGGATTGACTCACCCAGAAATTCTCCCTTTGTATGAGCTGTGAGA	1380
QY	1612	GAAGGGCTTGGGTCAACATCTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTG	1671
DB	1381	GAAGGGCTTGGGTCAACATCTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTG	1440
QY	1672	GTCGAAGGACGAGGGCTGTCTCCACAGAGCACGGGCCCTGTACTACCTGGGCCAGGCCCTG	1731
DB	1441	GTCGAAGGACGAGGGCTGTCTCCACAGAGCACGGGCCCTGTACTACCTGGGCCAGGCCCTG	1500
QY	1732	GAGGGTCTGGAATACCTCCACTCACGAAGGATTCTGCATGGGACGCTCAAAGCTGACAAC	1791
DB	1501	GAGGGTCTGGAATACCTCCACTCACGAAGGATTCTGCATGGGACGCTCAAAGCTGACAAC	1560
QY	1792	GTGCTCTGTCCAGCATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT	1851
DB	1561	GTGCTCTGTCCAGCATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTGT	1620
QY	1852	CTTCAACCTGATGGCTTGGGAAAGTCTTGCTCACAGGGGACTACATCCCTGGCACAGAG	1911
DB	1621	CTTCAACCTGATGGCTTGGGAAAGTCTTGCTCACAGGGGACTACATCCCTGGCACAGAG	1680
QY	1912	ACCCACATGGCTCGGAGTGGTGTCTGGGCGAGAGTGCACGCCAAGGTGGATGTCTGG	1971
DB	1681	ACCCACATGGCTCGGAGTGGTGTCTGGGCGAGAGTGCACGCCAAGGTGGATGTCTGG	1740
QY	1972	AGCAGCTGTGTATGATGCTGCACATGCTCAAAGGCTGCCACCCCTGGACTCAGTTCTTC	2031
DB	1741	AGCAGCTGTGTATGATGCTGCACATGCTCAAAGGCTGCCACCCCTGGACTCAGTTCTTC	1800
QY	2032	CGAGGGCGCTTGCCTCAAGATTGGCCAGCGAGCCTCCGGCTGTGAGGGAGATGCCACCC	2091
DB	1801	CGAGGGCGCTTGCCTCAAGATTGGCCAGCGAGCCTCCGGCTGTGAGGGAGATGCCACCC	1860
QY	2092	TCTTGCCGCCCTCTCACAGCCCAAGGCATCCACAGAGGGGTGAGGAAAAGGCCCATCCAC	2151
DB	1861	TCTTGCCGCCCTCTCACAGCCCAAGGCATCCACAGAGGGGTGAGGAAAAGGCCCATCCAC	1920
QY	2152	CGGCTGTCTCAGCGGAGCTGGAGGGAGGTGCAACCGGGCACTACAGCAAGTGGGAGGT	2211
DB	1921	CGGCTGTCTCAGCGGAGCTGGAGGGAGGTGCAACCGGGCACTACAGCAAGTGGGAGGT	1980
QY	2212	CTGAAGAGCCCTTTGGAGGGAGAAATAAAGAACCAAGACATCCACCGCCAAATCAAGCC	2271
DB	1981	CTGAAGAGCCCTTTGGAGGGAGAAATAAAGAACCAAGACATCCACCGCCAAATCAAGCC	2040
QY	2272	AATTACACAGACCTTCATGCCACCGCAGAGAGCTTTTCGCCAAGGGGCCCAAGGGCCC	2331
DB	2041	AATTACACAGACCTTCATGCCACCGCAGAGAGCTTTTCGCCAAGGGGCCCAAGGGCCC	2100

QY	2332	CGGCCAGCTGAGGAGACAACAGGCGAGAGCCCTTAAGCTCAGCCTCTCTCTCCACACAGAG	2399
DB	2101	CGGCCAGCTGAGGAGACAACAGGCGAGAGCCCTTAAGCTCAGCCTCTCTCTCCACACAGAG	2160
QY	2392	CCCCAGAGCCAAACAGTCTCTCTCCCTGACTTTTGGAGCAAGGAGGAGTCTGGATGTGG	2451
DB	2161	CCCCAGAGCCAAACAGTCTCTCTCCCTGACTTTTGGAGCAAGGAGGAGTCTGGATGTGG	2220
QY	2452	GAACCCCTTACCTCTGTCTCTCCCTGGAGCCAGCCCTTGCAGAAACCCAGCTCACCAGAG	2511
DB	2221	GAACCCCTTACCTCTGTCTCTCCCTGGAGCCAGCCCTTGCAGAAACCCAGCTCACCAGAG	2280
QY	2512	CGGAAAGCAACCGTCCCGGAGCAGGAACCTCAGCAGCTGGAAATAGAAATTTCTCTCAAC	2571
DB	2281	CGGAAAGCAACCGTCCCGGAGCAGGAACCTCAGCAGCTGGAAATAGAAATTTCTCTCAAC	2340
QY	2572	AGCCTGTCCCGACCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTGTGCCTTCAGCATC	2631
DB	2341	AGCCTGTCCCGACCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTGTGCCTTCAGCATC	2400
QY	2632	GACAGCCTCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAGAGCCTCTCAAAGCTCG	2691
DB	2401	GACAGCCTCTCCCTGTCCGATGACAGTGAGAGAACCCATCAAGAGCCTCTCAAAGCTCG	2460
QY	2692	CGGGACACCTGTGAGCTCAGGCGTACACTCTGTGAGCAGCACGAGCCGAGCTCGAAGCTCC	2751
DB	2461	CGGGACACCTGTGAGCTCAGGCGTACACTCTGTGAGCAGCACGAGCCGAGCTCGAAGCTCC	2520
QY	2752	AGCTGGAAACATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTG	2811
DB	2521	AGCTGGAAACATGTGTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTG	2580
QY	2812	GTGAAAGTCCAAATACAGTCTCTTAATGTGTGAACCTCGACATCCGGAGGTTCACCGG	2871
DB	2581	GTGAAAGTCCAAATACAGTCTCTTAATGTGTGAACCTCGACATCCGGAGGTTCACCGG	2640
QY	2872	GTCAAAGTGGGAGACATCGCACTGGCCATCAGCAGCCAGATCCCAAGCTCAGCCTTCAGC	2931
DB	2641	GTCAAAGTGGGAGACATCGCACTGGCCATCAGCAGCCAGATCCCAAGCTCAGCCTTCAGC	2700
QY	2932	TTGGTTCACCAAGACGGGCAGCCTGTTCCTACGACATGGAGGTTGCCAGACTCGGGCATC	2991
DB	2701	TTGGTTCACCAAGACGGGCAGCCTGTTCCTACGACATGGAGGTTGCCAGACTCGGGCATC	2760
QY	2992	GACCTGCAGTGCACACTGGCCCTGTATGTGCAGCTTCGCTCGAGCTGGAGGTTCAAGCAT	3051
DB	2761	GACCTGCAGTGCACACTGGCCCTGTATGTGCAGCTTCGCTCGAGCTGGAGGTTCAAGCAT	2820
QY	3052	GGCCAGCTGGAGAACAGGCCCTTAAACCTGCCTTCCACCGCCGGCTCCACCTGCCGAAA	3111
DB	2821	GGCCAGCTGGAGAACAGGCCCTTAAACCTGCCTTCCACCGCCGGCTCCACCTGCCGAAA	2880
QY	3112	GCAGCCTTCTCGTCTCGGTGCAGATGCTGCCCTTGAAAAACACAGGCTCAGCCGTTCCAGG	3171
DB	2881	GCAGCCTTCTCTGTCTCGGTGCAGATGCTGCCCTTGAAAAACACAGGCTCAGCCGTTCCAGG	2940
QY	3172	GGATTGCAGCCCCCGGCTTCAAGTGGGAACAGGGCTTCGACGACGAAGTGGGGGC	3231
DB	2941	GGATTGCAGCCCCCGGCTTCAAGTGGGAACAGGGCTTCGACGACGAAGTGGGGGC	3000
QY	3232	AAGCAGATGCTCCACAGATTTCAACCTGTAGAGCCTGCCACACCTGCTGAAAAACAT	3291
DB	3001	AAGCAGATGCTCCACAGATTTCAACCTGTAGAGCCTGCCACACCTGCTGAAAAACAT	3060
QY	3292	CGCCCACTGTAAGAGACAGAAGGAGGATGGCAGGAGTTACCTGGGGAAACAAAAACAGGA	3351
DB	3061	CGCCCACTGTAAGAGACAGAAGGAGGATGGCAGGAGTTACCTGGGGAAACAAAAACAGGA	3120
QY	3352	TCCTTTTCTGCCCCCTGCTCCAGTCCAGTTGGCCTGA	3387
DB	3121	TCCTTTTCTGCCCCCTGCTCCAGTCCAGTTGGCCTGA	3156

## RESULT 4

```

US-09-257-703-2
; Sequence 2, Application US/09257703
; Patent No. 6265538
; GENERAL INFORMATION:
; APPLICANT: Greene, Warner C.
; APPLICANT: Lin, Xin
; APPLICANT: Gelezuinas, Romas
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED
; TITLE OF INVENTION: BY TNF-ALPHA AND IL-1
; FILE REFERENCE: 30448.610U1
; CURRENT APPLICATION NUMBER: US/09/257,703
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: 60/076,299
; EARLIER FILING DATE: 1998-02-27
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2844
; TYPE: DNA
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)
US-09-257-703-2

```

Query Match	61.7%;	Score 2837.6;	DB 3;	Length 2844;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2840;	Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;
QY	232	ATGGCAGTGTGAAATGGCTGCCCAGGTGCCCCCTGGCTCAGCAGATGGGCGACAGAA	291	
DB	1	ATGGCAGTGTGAAATGGCTGCCCAGGTGCCCCCTGGCTCAGCAGATGGGCGACAGAA	60	
QY	292	GAATCTCCCAAGCCAAAGGAGAGACGCCGCCACTGGGGGAGAAACAGAGCTCCGCTTAC	351	
DB	61	GAATCTCCCAAGCCAAAGGAGAGACGCCGCCACTGGGGGAGAAACAGAGCTCCGCTTAC	120	
QY	352	AAGCTTTGAGCCCGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTCAATGAC	411	
DB	121	AAGCTTTGAGCCCGTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTCAATGAC	180	
QY	412	GTGATTACCAAGGGCA CAGCCAAGGAAGGCTCCGAGGACGGCCAGCTGCGCATCTCTATC	471	
DB	181	GTGATTACCAAGGGCA CAGCCAAGGAAGGCTCCGAGGACGGCCAGCTGCGCATCTCTATC	240	
QY	472	ATCGCCCGAGCTGAGTGTGAGATAGCAAGAGTTTCAGCCGCCACCTTTTCAGAACGCAAT	531	
DB	241	ATCGCCCGAGCTGAGTGTGAGATAGCAAGAGTTTCAGCCGCCACCTTTTCAGAACGCAAT	300	
QY	532	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT	591	
DB	301	TTTCATCGCTGGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAT	360	
QY	592	GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAAGGAAAGCGTCGAGC	651	
DB	361	GTGGCCCATCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAAGGAAAGCGTCGAGC	420	
QY	652	AAAGCCCGGAAGAAACGGAAAGAAAGAGTCAAAGTCCCTGGCTCATGACGAGTGGCC	711	
DB	421	AAAGCCCGGAAGAAACGGAAAGAAAGAGTCAAAGTCCCTGGCTCATGACGAGTGGCC	480	
QY	712	TTGGCCAAACCCCTCCCAAGACCCCTGAGCAGAGAGCTGCACCATCCCAAGTCAGAG	771	
DB	481	TTGGCCAAACCCCTCCCAAGACCCCTGAGCAGAGAGCTGCACCATCCCAAGTCAGAG	540	
QY	772	GATGAGTCTCCACTCGCGGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCCTTG	831	
DB	541	GATGAGTCTCCACTCGCGGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCCCTTG	600	
QY	832	AAGAAACAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGCGCGGCT	891	
DB	601	AAGAAACAGGCTTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTACGCGCGGCT	660	
QY	892	CTGCGCTCGATCAGAACTCCAAACATGATCAGCCCTTTGCAATGTCTGAACACAGCTGG	951	

QY	2032	CGAGGCCGCTCTGCTC	CAAGATTGCGACGAGCCTCGCGCTGTGAGGGAGATCCCAACC	2099
DB	1801	CGAGGCCGCTCTGCTC	CAAGATTGCGACGAGCCTCGCGCTGTGAGGGAGATCCCAACC	1860
QY	2092	TCCTGCGCCCTCTC	CACAGGCCACAGSCCATCCAAAGAGGGCTGAGAAAGAGCCCATCCAC	2151
DB	1861	TCCTGCGCCCTCTC	CACAGGCCACAGSCCATCCAAAGAGGGCTGAGAAAGAGCCCATCCAC	1920
QY	2152	CGCGTGTCTGACGGAGCT	GGGAGGGAAAGTGAAACCGGGCACTACAGCAAGTGGGAGGT	2211
DB	1921	CGCGTGTCTGACGGAGCT	GGGAGGGAAAGTGAAACCGGGCACTACAGCAAGTGGGAGGT	1980
QY	2212	CTGAAGAGCCCTTGAGGGGAGAA	TATAAGAAACCAAGACATCCACCGCCAAATCAAGCC	2271
DB	1981	CTGAAGAGCCCTTGAGGGGAGAA	TATAAGAAACCAAGACATCCACCGCCAAATCAAGCC	2040
QY	2272	AATTACCACAGACCTTCAT	TGCCCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCAAGGGCCC	2331
DB	2041	AATTACCACAGACCTTCAT	TGCCCCAGCCGAGAGAGCTTTCCGCAAGGGGCCCAAGGGCCC	2100
QY	2332	CGGCCAGCTGAGGAGACA	AAAGGAGAGCCCTTAAGCTCAGCCTCTCTCCCAACCAAG	2391
DB	2101	CGGCCAGCTGAGGAGACA	AAAGGAGAGCCCTTAAGCTCAGCCTCTCTCCCAACCAAG	2160
QY	2392	CCCCAGAGCCAAACAAGTCT	CCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGG	2451
DB	2161	CCCCAGAGCCAAACAAGTCT	CCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGG	2220
QY	2452	GAAACCTTTACTCTGTCT	CTCCTCGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG	2511
DB	2221	GAAACCTTTACTCTGTCT	CTCCTCGGAGCCAGCCCTGCCAGAAACCCAGCTCACCAGAG	2280
QY	2512	CGGAAAGCAACCGTCCGG	AGCAGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAAC	2571
DB	2281	CGGAAAGCAACCGTCCGG	AGCAGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAAC	2340
QY	2572	AGCCTGTCCCAGCCATT	TTCTCTGGAGGACAGGAGCAAAATTTCTCGTGCCTCAGCATC	2631
DB	2341	AGCCTGTCCCAGCCATT	TTCTCTGGAGGACAGGAGCAAAATTTCTCGTGCCTCAGCATC	2400
QY	2632	GACAGCTCTCCCTGTG	CGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGGCTCG	2691
DB	2401	GACAGCTCTCCCTGTG	CGATGACAGTGAAGAAACCCATCAAGGCTCTCAAGGCTCG	2460
QY	2692	CGGACACCCCTGAGCT	CAGGCGTACACTCTGGAGACGACGAGCGTCCAGAGCTCC	2751
DB	2461	CGGACACCCCTGAGCT	CAGGCGTACACTCTGGAGACGACGAGCGTCCAGAGCTCC	2520
QY	2752	AGCTGGAAACATG	GTGGTGGCCCGGGGCGGCCCAACGACACCCCAAGCTATTTCAATGGT	2811
DB	2521	AGCTGGAAACATG	GTGGTGGCCCGGGGCGGCCCAACGACACCCCAAGCTATTTCAATGGT	2580
QY	2812	GTGAAAGTCCAAATAC	AGTCTCTTAATGTGTAACCTGACATCCGGGAGTTCCACCGG	2871
DB	2581	GTGAAAGTCCAAATAC	AGTCTCTTAATGTGTAACCTGACATCCGGGAGTTCCACCGG	2640
QY	2872	GTCAAAGTGGGAGACAT	TCGCACCTGGCATCAGACGCCAGATCCACGCTGCAGGCTTTCAGC	2931
DB	2641	GTCAAAGTGGGAGACAT	TCGCACCTGGCATCAGACGCCAGATCCACGCTGCAGGCTTTCAGC	2700
QY	2932	TTGGTACCAAAGACGG	GCACCTGTTTCGCTTACGACATGAGAGTGCAGACTCGGGCATC	2991
DB	2701	TTGGTACCAAAGACGG	GCACCTGTTTCGCTTACGACATGAGAGTGCAGACTCGGGCATC	2760
QY	2992	GACCTGCAGTGCAC	ACTGAGTGGCCCTTGATGGCAGCTTCGCTGAGAGCTGGAGGTCAGCAT	3051
DB	2761	GACCTGCAGTGCAC	ACTGAGTGGCCCTTGATGGCAGCTTCGCTGAGAGCTGGAGGTCAGCAT	2820
QY	3052	GGCCAGCTGGAGAAC	AGGCCCTTAA 3075	
DB	2821	GGCCAGCTGGAGAAC	AGGCCCTTAA 2844	

RESULT 5  
US-09-871-889A-2  
; Sequence 2, Application US/09871889A  
; Patent No. 6645728  
; GENERAL INFORMATION:  
; APPLICANT: Greene, Warner C.  
; APPLICANT: Lin, Xin  
; APPLICANT: Gelezuinas, Romas  
; TITLE OF INVENTION: A NOVEL INHIBITOR OF THE INFLAMMATORY RESPONSE INDUCED BY TNF-ALPHA  
; TITLE OF INVENTION: IL-1  
; FILE REFERENCE: 30448.61USD1  
; CURRENT APPLICATION NUMBER: US/09/871,889A  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 09/257,703  
; PRIOR FILING DATE: 1999-02-25  
; PRIOR APPLICATION NUMBER: 60/076,299  
; PRIOR FILING DATE: 1998-02-27  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 2844  
; TYPE: DNA  
; ORGANISM: NF-KAPPA B INDUCING KINASE (NIK)  
US-09-871-889A-2

Query Match 61.7%; Score 2837.6; DB 4; Length 2844;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2840; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

2

Qy 892 CTGCTCGATCAGAACTCCAACTGATCAGCCCCCTTGCAATGTCTGAACCAACGCTGTGG 951  
Db 661 CTGCTCGATCAGAACTCCAACTGATCAGCCCCCTTGCAATGTCTGAACCAACGCTGTGG 720  
Qy 952 AAATCTGACACACCCAGGACGGAGGCCCTGCGCCCTGCCCCAGCACCCCTTCCCTAT 1011  
Db 721 AAATCTGACACACCCAGGACGGAGGCCCTGCGCCCTGCCCCAGCACCCCTTCCCTAT 780  
Qy 1012 AGCAGATGCTCATCCCTTCCCATTCACACCTCTCAGACCCCTGGAACCTTCACCCCTGT 1071  
Db 781 AGCAGATGCTCATCCCTTCCCATTCACACCTCTCAGACCCCTGGAACCTTCACCCCTGT 840  
Qy 1072 GAGTCTTCTGGGAAACTCGCTGTGTAGACAGCCAGAAACCTTCTGCTGACCCACAC 1131  
Db 841 GAGTCTTCTGGGAAACTCGCTGTGTAGACAGCCAGAAACCTTCTGCTGACCCACAC 900  
Qy 1132 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCC 1191  
Db 901 CTGAGCAAACTGGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCC 960  
Qy 1192 AGCTGCTGTCTCGTGTGGCCCATGAGAAATTTCTGTGAGGAATACCTAGTGCATGCT 1251  
Db 961 AGCTGCTGTCTCGTGTGGCCCATGAGAAATTTCTGTGAGGAATACCTAGTGCATGCT 1020  
Qy 1252 CTGCAAGCAGCGTGAGCTCAAGCCAGGCCACAGCCCTGACAGCCTGGCCAAAGCCTGG 1311  
Db 1021 CTGCAAGCAGCGTGAGCTCAAGCCAGGCCACAGCCCTGACAGCCTGGCCAAAGCCTGG 1080  
Qy 1312 GCAGACGGGCTCAGATCCCGGGAGCCAGCCCAAACTGAGGACCAAGAGGTGTC 1371  
Db 1081 GCAGACGGGCTCAGATCCCGGGAGCCAGCCCAAACTGAGGACCAAGAGGTGTC 1140  
Qy 1372 CTGCTCACTGAGAACTCAAGCCAGTGTATGATACCGAGAAAGTCCACTGGGCC 1431  
Db 1141 CTGCTCACTGAGAACTCAAGCCAGTGTATGATACCGAGAAAGTCCACTGGGCC 1200  
Qy 1432 ACGCACAGCTCCGCTGGGAGAGGCTCTTCGGAGAGGTGCACAGGATGGAGCAAG 1491  
Db 1201 ACGCACAGCTCCGCTGGGAGAGGCTCTTCGGAGAGGTGCACAGGATGGAGCAAG 1260  
Qy 1492 CAGACTGCTTCAGTGGCTGTCAAAAGGTGGCTGCGCTGGAAGTATTTGGGGCAGAGAG 1551  
Db 1261 CAGACTGCTTCAGTGGCTGTCAAAAGGTGGCTGCGCTGGAAGTATTTGGGGCAGAGAG 1320  
Qy 1552 CTGATGCACTGTGAGAGTTGACCTCACCCAGAAATGTCCCTTTGATGGAGCTGTGAGA 1611  
Db 1321 CTGATGCACTGTGAGAGTTGACCTCACCCAGAAATGTCCCTTTGATGGAGCTGTGAGA 1380  
Qy 1612 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1671  
Db 1381 GAAGGGCTTGGGTCAACATCTTCATGGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTG 1440  
Qy 1672 GTCAAGGAGCAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACTTGGGGCAGGCCCTG 1731  
Db 1441 GTCAAGGAGCAGGGCTGTCTCCAGAGAGACCGGGCCCTGTACTACTTGGGGCAGGCCCTG 1500  
Qy 1732 GAGGCTGTGGAATACCTCCACTCAGAGGATTCGCAATGGGAGCTCAAGCTGCAAC 1791  
Db 1501 GAGGCTGTGGAATACCTCCACTCAGAGGATTCGCAATGGGAGCTCAAGCTGCAAC 1560  
Qy 1792 GTGCTCTGTCCAGCGATGGAGCCACGACGCCCTCTGTGATTTGGCCATGTGTGTGT 1851  
Db 1561 GTGCTCTGTCCAGCGATGGAGCCACGACGCCCTCTGTGATTTGGCCATGTGTGTGT 1620  
Qy 1852 CTTCAACTGATGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCCCTGGCACAGAG 1911  
Db 1621 CTTCAACTGATGCTGGGAAAGTCTTTGCTCAAGGGGACTACATCCCTGGCACAGAG 1680  
Qy 1912 ACCACATGGCTCCGAGGTGGTCTGGGAGGAGCTGCGACGCCAAGGTGGATGCTGG 1971  
Db 1681 ACCACATGGCTCCGAGGTGGTCTGGGAGGAGCTGCGACGCCAAGGTGGATGCTGG 1740

Qy 1972 AGCAGCTGTGTATGATGTGCACATGCTCAAAGGCTGGCACCCCTGAGACTCAGTTCTTTC 2031  
Db 1741 AGCAGCTGTGTATGATGTGCACATGCTCAAAGGCTGGCACCCCTGAGACTCAGTTCTTTC 1800  
Qy 2032 CGAGGGCCGCTCTGCTCCTCAAGATTTGCCAGAGGCTCCGCTGTGAGGGAGATCCACCC 2091  
Db 1801 CGAGGGCCGCTCTGCTCCTCAAGATTTGCCAGAGGCTCCGCTGTGAGGGAGATCCACCC 1860  
Qy 2092 TCTTGGCCCTCTCTCAGAGCCAGGCCATCAAAGAGGGGCTGAGGAAAGAGCCCATCCAC 2151  
Db 1861 TCTTGGCCCTCTCTCAGAGCCAGGCCATCAAAGAGGGGCTGAGGAAAGAGCCCATCCAC 1920  
Qy 2152 CGCTGTCTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGGACTACAGCAAGTGGAGGT 2211  
Db 1921 CGCTGTCTGAGCGGAGCTGGGAGGGAAGGTGAACCGGGGACTACAGCAAGTGGAGGT 1980  
Qy 2212 CTGAAGAGCCCTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2271  
Db 1981 CTGAAGAGCCCTTGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCC 2040  
Qy 2272 AATTACCAACAGACCTCTCATGCCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCC 2331  
Db 2041 AATTACCAACAGACCTCTCATGCCCAGCCGAGAGAGCTTTGCGCAAGGGCCCCAGGGCCC 2100  
Qy 2332 CGCCAGCTGAGGAGACAAAGCAGAGAGCCCTTAAGCTCAGACCTCTCTCCACACAGAG 2391  
Db 2101 CGCCAGCTGAGGAGACAAAGCAGAGAGCCCTTAAGCTCAGACCTCTCTCCACACAGAG 2160  
Qy 2392 CCCCAGAGCCAAACAAAGTCTCTCCCTGACCTTTGAGCAAGGAGGAGTCTGGGATGTGG 2451  
Db 2161 CCCCAGAGCCAAACAAAGTCTCTCCCTGACCTTTGAGCAAGGAGGAGTCTGGGATGTGG 2220  
Qy 2452 GAACCCCTTACTCTCTCTCTGAGCCAGCCCTGCGAGAGAGCTCCAGAAACCCAGCTCACCAGAG 2511  
Db 2221 GAACCCCTTACTCTCTCTCTGAGCCAGCCCTGCGAGAGAGCTCCAGAAACCCAGCTCACCAGAG 2280  
Qy 2512 CGGAAGCAACCGTCCCGAGCAGGAACTGCGAGCAGCTGGAATAAGATTTATTCCTCAAC 2571  
Db 2281 CGGAAGCAACCGTCCCGAGCAGGAACTGCGAGCAGCTGGAATAAGATTTATTCCTCAAC 2340  
Qy 2572 AGCTGTCCAGCCATTTCTCTGAGGAGCAGGACAAATTTCTCTGCTGCTCAGCATC 2631  
Db 2341 AGCTGTCCAGCCATTTCTCTGAGGAGCAGGACAAATTTCTCTGCTGCTCAGCATC 2400  
Qy 2632 GACAGCTCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAAGGCTCTCAAAGCTCG 2691  
Db 2401 GACAGCTCTCCCTGTCCGATGACAGTGAAGAAACCCATCAAAGGCTCTCAAAGCTCG 2460  
Qy 2692 CGGAGACCTGAGCTCAGCGCTACACTCTCTGAGCAGCAGCGCGAGGCTCGAAGTCC 2751  
Db 2461 CGGAGACCTGAGCTCAGCGCTACACTCTCTGAGCAGCAGCGCGAGGCTCGAAGTCC 2520  
Qy 2752 AGCTGGAACATGTGTGGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGT 2811  
Db 2521 AGCTGGAACATGTGTGGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGT 2580  
Qy 2812 GTGAAGTCCAAATACAGTCTCTTAATGTGTGAACCTGTGCACATCCGGAGTTCACCGG 2871  
Db 2581 GTGAAGTCCAAATACAGTCTCTTAATGTGTGAACCTGTGCACATCCGGAGTTCACCGG 2640  
Qy 2872 GTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGACGCTTCAGC 2931  
Db 2641 GTCAAGTGGGAGACATCGCCACTGGCATCAGCAGCAGATCCAGCTGACGCTTCAGC 2700  
Qy 2932 TTGCTCAACAAAGCGGGCAGCTGTTCGCTACGACATGAGGAGTCCAGACTCGGGCATC 2991  
Db 2701 TTGCTCAACAAAGCGGGCAGCTGTTCGCTACGACATGAGGAGTCCAGACTCGGGCATC 2760  
Qy 2992 GACCTCAGTGCACACTGGGCCCTGTATGGCAGCTTCGCTTGGAGCTGAGGGTCAAGCAT 3051  
Db 2761 GACCTCAGTGCACACTGGGCCCTGTATGGCAGCTTCGCTTGGAGCTGAGGGTCAAGCAT 2820  
Qy 3052 GGCCAGCTGGAGAACAGGCCCTTAA 3075

```
Db 2821 GGCACGCTGGAGAACAGGCCCTTAA 2844
|||||
RESULT 6
US-09-513-999C-2583
; Sequence 2583, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2583
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 79..360
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 19
; OTHER INFORMATION: s=g or c
; NAME/KEY: misc_feature
; LOCATION: 22
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-2583

Query Match 6.5%; Score 298.2; DB 4; Length 362;
Best Local Similarity 95.9%; Pred. No. 3.2e-70;
Matches 306; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 194 GTGTTTGTTCTCTCAGGATGAGCACAGCCCTGGGAGATGGCAGTGATGGAAATGGCCT 253
Db 41 GAGATCGGGCGCTGGCCGATGAGCACAGCCCTGGGAGATGCGATGGAAATGGCCT 100
QY 254 GCCCAGGTCCCTGGCTCAGCAGTGGGCGACAGAGAACTCCCAAGCCAAAGGAGA 313
Db 101 GCCCAGGTCCCTGGCTCAGCAGTGGGCGACAGAGAACTCCCAAGCCAAAGGAGA 160
QY 314 AGACGCCGCCACTGGGGAAGAACAGAGCTCCGCTCTACAAGCTTCAGGCCGTGGAGAAGA 373
Db 161 AGACGCCGCCACTGGGGAAGAACAGAGCTCCGCTCTACAAGCTTCAGGCCGTGGAGAAGA 220
QY 374 GCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAATGAGCTGATTACCAAGGGCAGAGCA 433
Db 221 GCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAATGAGCTGATTACCAAGGGCAGAGCA 280
QY 434 AGAAGGCTCCAGGAGGCCGAGCTGCCATCTCTATATGCCAGGCTGAGTGTGAGA 493
Db 281 AGAAGGCTCCAGGAGGCCGAGCTGCCATCTCTATATGCCAGGCTGAGTGTGAGA 340
QY 494 ATAGCCAGAGTTTCAGCCC 512
Db 341 ATAGCCAGAGTTTCAGCCC 359

RESULT 7
US-09-949-016-5071
; Sequence 5071, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5071
; LENGTH: 2207
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5071

Query Match 1.9%; Score 88.8; DB 4; Length 2207;
Best Local Similarity 53.4%; Pred. No. 2.8e-13;
Matches 236; Conservative 0; Mismatches 197; Indels 9; Gaps 2;

QY 1625 TCACATCTTCATGGAGCTGCTGGAAGCTGCTCCCTGGCCAGCTGCTCAAGGAGCAGG 1684
Db 1312 TGACCATCTTCATGGAGTACATGCCAGGGGCTCGTGAAGACACAGTTGAAGGCTTAGC 1371
QY 1685 GCTGTCCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGCCCTGGAGGCTCTGGAAT 1744
Db 1372 GTGCTCTGACAGAGAGCGTGACCCGAAAGTACACGCGCAGATCCTGGAGGCGATGCTCT 1431
QY 1745 ACCTCCACTCAGGAAGATTCTGCATGGGAGCTCAAGCTGACAACTGCTCTCTGTCCA 1804
Db 1432 ACCTGCACAGCAACATGATTGTTACCGGGGACATTAAGGGGAGCCAACTCTCCGAGACT 1491
QY 1805 GCGATGGGAGCGACGACGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACTGATG 1864
Db 1492 CTGCTGGGA---ATGTAAGCTGGGGGACTTTGGGCCAGCAAGCGCTGCAGAC----- 1543
QY 1865 GCCTGGGAAAGTCTTGTCTACAGGGGACTACATCCCTGGCAGACAGACCACATGGGCTC 1924
Db 1544 -GATCTGTATGTCGGGACGGGCGATGCGCTCCGTCACCTGGCACACCCCTACTGGATGAGCC 1602
QY 1925 CGGAGGTGCTGCTGGGAGGAGCTGCGACGCCAAGGTGATGCTGGAGCAGCTGCTGTA 1984
Db 1603 CTGAGGTGATCAGCGCGCAGGGGCTATGGAAGGAAGCAGACGCTGTGGAGCCTGGGCTGCA 1662
QY 1985 TGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCGCTCT 2044
Db 1663 CTGTGGTGGAGTCTGACAGAGAAACACCGTGGCAGAGTATGAAGCTATGAGCTATGCGCGCA 1722
QY 2045 GCCTCAAGATTGCCAGGAGCC 2066
Db 1723 TCTTCAAGATTGCCACCCAGCC 1744

RESULT 8
US-09-658-688A-3
; Sequence 3, Application US/09658688A
; Patent No. 6498035
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: William Gaarde
; APPLICANT: Brett P. Monia
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF MEK3 EXPRESSION
; FILE REFERENCE: RTS-0143
; CURRENT APPLICATION NUMBER: US/09/658,688A
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 3
; LENGTH: 2348
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (84)...(1964)
US-09-658-688A-3

Query Match      1.9%; Score 88.8; DB 4; Length 2348;
Best Local Similarity 53.4%; Pred. No. 2.9e-13;
Matches 236; Conservative 0; Mismatches 197; Indels 9; Gaps 2;

Qy 1625 TCAACATCTTCATGGAGTCTGGAAGGTGGCTCCCTGGGGCAGCTGGTCAAGGAGCAGG 1684
Db 1399 TGACCACTTCATGGAGTACATGCGAGGGGCTCGTGAAGACAGATTGAGGCTTACG 1458

Qy 1685 GCTGTCTCCAGAGGACGGGCCCTGTACTACCTGGGCGAGGCCCTGGAGGGTCTGGAAT 1744
Db 1459 GTGCTCTGACAGAGCGTGAACCCGAAAGTACACGCGCAGATCCTGGAGGGCATGTCCT 1518

Qy 1745 ACCTCCACTCAGAGAGATTCTGATGGGAGCTCAAGAGCTGACAAGCTGCTCTGTCCA 1804
Db 1519 ACCTGCAAGCAACATGATTGTTACCGGGACATTAAGGGAGCCAAACATCCTCCGAGACT 1578

Qy 1805 GCGATGGAGCACGCGAGCCCTGTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864
Db 1579 CTGCTGGGA---ATGTAAGCTGGGGACTTTGGGGCCAGCAAGCGCTGCAGAC----- 1630

Qy 1865 GCCTGGGAAAGTCTTGTCTCAAGGGGACTACATCCCTGGCAGAGACCCACATGGCTC 1924
Db 1631 -GATCTGTATGTCGGGAGCGGCGTCCGCTCCGTCACCTGGCAGACCCCTACTGGATGAGCC 1689

Qy 1925 CGGAGGTGGTCTGGGAGGAGCTGGCAGCCAGGCTGGATGCTTGGAGCAGCTGCTGTGA 1984
Db 1690 CTGAGGTGATCAGCGGCGAGGGCTATGGAAGAAAGCAGACGCTGGAGCGCTGGGCTGCA 1749

Qy 1985 TGATGCTGCACATGCTCAAGCGCTGCCACCCCTGCAGTCTTCTCCGAGGGCGGCTCT 2044
Db 1750 CTGTGGTGGAGTGTGACAGAGAAACACCGTGGGAGAGTATGAAGCTATATGGCCGCA 1809

Qy 2045 GCCTCAAGATTGCCAGCGAGCC 2066
Db 1810 TCTTCAAGATTGCCACCCAGCC 1831

RESULT 9
US-08-472-934-5
; Sequence 5, Application US/08472934
; Patent No. 5753446
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, GARY L.
; TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
; TITLE OF INVENTION: RESPONSIVENESS TO EXTERNAL SIGNALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lnhive and Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,934
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/440,421
; FILING DATE: 15-May-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/354,516
```

```
; FILING DATE: 21-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 5,405,941
; FILING DATE: 15-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/323,460
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11690
; FILING DATE: 14-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04178
; FILING DATE: 15-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Giulio A. DeConti, Jr. Esq.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: CPI-004DVC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3089 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 400..2280
US-08-472-934-5

Query Match      1.8%; Score 83.8; DB 1; Length 3089;
Best Local Similarity 52.6%; Pred. No. 7.4e-12;
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;

Qy 1625 TCAACATCTTCATGGAGTCTGGAAGGTGGCTCCCTGGGGCAGCTGGTCAAGGAGCAGG 1684
Db 1715 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAAGACCAAGTTGAAGGCCCTACG 1774

Qy 1685 GCTGTCTCCAGAGGACGGGCCCTGTACTACCTGGGCGAGGCCCTGGAGGGTCTGGAAT 1744
Db 1775 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGCGAGATTCTGGAGGGCATGTCTAT 1834

Qy 1745 ACCTCCACTCAGCAAGGATTCTGCATGGGAGCGTCAAGAGCTGACAACGCTGCTCTGTCCA 1804
Db 1835 ACCTGACAGCAACATGATTGTGATCGGACATCAAGGGAGCCAAATATCTCTCCGAGACT 1894

Qy 1805 GCGATGGAGCGCAGCAGCCCTCTGTGACTTTTGGCCATGCTGTGTGTCTTCAACCTGATG 1864
Db 1895 CAGCTGGGA---ATGTGAAGCTTGGGGATTTTGGGGCCAGCAACCGCTACAGACCATCT 1951

Qy 1865 GCCTGGGAAAGTCTTGTCTCACAGGGGACTACATCCCTGGCAGAGACCCACATGGCTC 1924
Db 1952 GCATGTGAGGACAGGCAAT-----TCGCTCTGTCTACTGGCAGACACCTACTTGGATGAGTC 2005

Qy 1925 CGGAGGTGGTCTGGGCGAGGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGTA 1984
Db 2006 CTGAAGTCATCAGTGGCGAGGGCTATGGAAGAAAGGAGCAGCGTGTGGAGCGCTGGGCTGTA 2065

Qy 1985 TGATGTGACATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTCTCGAGGGCGGCTCT 2044
Db 2066 CTGTGTGGAAATGCTGACAGAGAAACCACTTTGGGCGAGAGTATGAAGCTATATGGCTGCA 2125

Qy 2045 GCCTCAAGATTGCCAGCGAGCCCTCC 2069
Db 2126 TTTTCAAGATTGCCACCAGCCCTAC 2150

RESULT 10
US-08-323-460A-5
; Sequence 5, Application US/08323460A
; Patent No. 5854043
```







Qy 1695 GCTGCTCCAGAGACCGGCGCTGTACTACCTGGGCGAGCCCTGGAGGGTCTGAAT 1744  
Db 1775 GAGCTCTGACAGAGAGTGTGACCGCAAGTACACCCCGCAGATTCTGGAGGGCATGTCT 1834  
Qy 1745 ACCTCCACTCAGGAAGATTCTGATGGGACGTCAGAACTGACAACTGCTCTGCTGCA 1804  
Db 1835 ACCTGCACAGCAACATGATTGTGATCGGACATCAGGGAGCCATATCTCCGAGACT 1894  
Qy 1805 GCGATGGAGCCAGCGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864  
Db 1895 CAGCTGGGA---ATGTGAAGCTTGGGATTTGGGGCCAGCAACGCCCTACAGACCATCT 1951  
Qy 1865 GCCTGGGNAAGTCTTGTCTACAGGGGACTACATCCCTGGCAGACAGACCCACATGGCTC 1924  
Db 1952 GCATGTCCAGGACAGGCAT-----TCGCTGTGCTCAGTGGCAGCACCCTACTGGATGATC 2005  
Qy 1925 CGGAGGTGGTCTGGGAGGAGCTGCGACGCCAAGGTGATGCTGGAGCAGCTGCTGTA 1984  
Db 2006 CTGAAGTCATCAGTGGCGAGGGCTATGGAAGAGGCGAGCTGTGAGCGCTGGGCTGTA 2065  
Qy 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTCT 2044  
Db 2066 CTGTGGTGGAAATGCTGACAGAGAAACCACTTGGGCGAGATATGAAGCTATGCTGCCA 2125  
Qy 2045 GCCTCAAGATTGCCAGCGAGCCTCC 2069  
Db 2126 TTTTCAAGATTGCCACCCAGCCTAC 2150

RESULT 12

US-08-461-145C-5  
; Sequence 5, Application US/08461145C  
; Patent No. 6074861  
; GENERAL INFORMATION:  
; APPLICANT: JOHNSON, GARY L.  
; TITLE OF INVENTION: NOVEL MEKK PROTEINS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lahive and Cockfield, LLP  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,145C  
; FILING DATE: 5-JUNE-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/049,254  
; FILING DATE: 11-APR-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/323,460  
; FILING DATE: 14-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/11690  
; FILING DATE: 14-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/04178  
; FILING DATE: 15-APR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/354,516  
; FILING DATE: 21-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kara, Catherine J.  
; REGISTRATION NUMBER: P-41,106  
; REFERENCE/DOCKET NUMBER: CPI-004CN1  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3089 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 400..2280  
; US-08-461-145C-5

Query Match 1.8%; Score 83.8; DB 3; Length 3089;  
Best Local Similarity 52.6%; Pred. No. 7,4e-12;  
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;  
Qy 1625 TCAACATCTTCATGGAGCTGCTGGAAGTGGCTCCCTGGGCGAGCTGGTCAAGAGCAGG 1684  
Db 1715 TCACCATCTTTATGGAGTATATGCCAGGGGCTCTGTAAAGACCCAGTTGAAGGCCCTACG 1774  
Qy 1685 GCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGCCAGGCCCTGGAGGCTCTGGAAT 1744  
Db 1775 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGCGAGATTCTGGAGGCGATGTCT 1834  
Qy 1745 ACCTCCACTCAGCAAGGATTCTGCATGGGAGCTGCAAAAGCTGACAAAGCTGCTCTGCTCA 1804  
Db 1835 ACCTGCACAGCAACATGATTGTGATCGGACATCAAGGGAGCCATATCTCCGAGACT 1894  
Qy 1805 GCGATGGAGCCAGCGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864  
Db 1895 CAGCTGGGA---ATGTGAAGCTTGGGATTTGGGGCCAGCAACGCCCTACAGACCATCT 1951  
Qy 1865 GCCTGGGAAAGTCTTGTCTCACAGGGGACTACATCCCTGGCAGAGACCCACATGGCTC 1924  
Db 1952 GCATGTCAGGACAGGCAT-----TCGCTGTCTACTGGCAGACCCCTACTGGATGATC 2005  
Qy 1925 CGGAGGTGGTCTGGGCGAGGAGCTGCGACGCCAAGGTGATGCTGGAGCAGCTGCTGTA 1984  
Db 2006 CTGAAGTCATCAGTGGCGAGGGCTATGGAAGAGGCGAGCGTGTGGAGCCTGGGCTGTA 2065  
Qy 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCGCTCT 2044  
Db 2066 CTGTGGTGGAAATGCTGACAGAGAAACCACTTGGGCGAGATATGAAGCTATGCTGCCA 2125  
Qy 2045 GCCTCAAGATTGCCAGCGAGCCTCC 2069  
Db 2126 TTTTCAAGATTGCCACCCAGCCTAC 2150

RESULT 13

US-08-628-829-9  
; Sequence 9, Application US/08628829A  
; Patent No. 633170  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Gary L.  
; TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To External  
; FILE REFERENCE: CPI-004DVP3  
; CURRENT APPLICATION NUMBER: US/08/628,829A  
; FILING DATE: 1996-04-05  
; EARLIER APPLICATION NUMBER: 08/440,421  
; EARLIER FILING DATE: 1995-05-15  
; EARLIER APPLICATION NUMBER: 08/323,460  
; EARLIER FILING DATE: 1994-10-14  
; EARLIER APPLICATION NUMBER: 08/049,254  
; EARLIER FILING DATE: 1993-05-15  
; EARLIER APPLICATION NUMBER: 08/410,602  
; EARLIER FILING DATE: 1995-04-24  
; EARLIER APPLICATION NUMBER: 08/472,934  
; EARLIER FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9  
; LENGTH: 3089  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (400)..(2277)  
US-08-628-829-9

Query Match  
Best Local Similarity 1.8%; Score 83.8; DB 3; Length 3089;  
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;  
QY 1625 TCAACATCTTTCAGAGCTGCTGGAAGGTGGCTCCCTGGCCAGCTGGTCAAGGACAGG 1684  
Db 1715 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAGACCAAGTTGAAGGCTACG 1774  
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGCCCTGGAGGGTCTGGAAT 1744  
Db 1775 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGGCAGATTTCTGGAGGGCATGTCTAT 1834  
QY 1745 ACCTCCACTCACGAAGGATTTGCAATGGGAGCTCAAGCTGACACGCTGCTCCTGTCCA 1804  
Db 1835 ACCTGCACAGCAACATGATTTGCAATGGGAGCATCAAGGAGGCCAATATCTCCGAGACT 1894  
QY 1805 GCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864  
Db 1895 CAGCTGGGA--ATGTGAAGCTTGGGATTTTGGGCCAGCAACGCCCTAGACCATCT 1951  
QY 1865 GCCTGGGAAGTCTTGTCTACAGGGGACTATACCTCCCTGGCAAGAGACCCACATGGCTC 1924  
Db 1952 GCATGTCCAGGACAGGCAT-----TCGCTCTGTCACTGGCACACCCCTACTGGATGATC 2005  
QY 1925 CGAGGTGTGTGGGCGAGAGCTGCGACGCCAAGGTGATGCTTCCGAGGCGGCTCT 2044  
Db 2066 CTGTGTGGAATGCTGACAGAGAACCCACTTGGGCGAGATGAGCTATGAGCTATGGCTGCTGA 2150  
QY 2045 GCCTCAAGATTGCCAGGAGCCCTCC 2069  
Db 2126 TTTTCAAGATTGCCAGCCAGCCTAC 2150

RESULT 14  
US-09-423-890-11  
; Sequence 11, Application US/09423890  
; Patent No. 6312934  
; GENERAL INFORMATION:  
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; FILE REFERENCE: CPI-085CPPC  
; CURRENT APPLICATION NUMBER: US/09/423,890  
; CURRENT FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: USN 60/078,153  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: USN 60/099,165  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 3332  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (332)..(2209)  
US-09-423-890-11

Query Match  
Best Local Similarity 1.8%; Score 83.8; DB 3; Length 3332;  
Matches 233; Conservative 0; Mismatches 203; Indels 9; Gaps 2;

Best Local Similarity 52.6%; Pred. No. 7.7e-12;  
Matches 234; Conservative 0; Mismatches 202; Indels 9; Gaps 2;  
QY 1625 TCAACATCTTTCAGAGCTGCTGGAAGGTGGCTCCCTGGGCGAGCTGGTCAAGGACAGG 1684  
Db 1647 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAAGACCAAGTTGAAGGCTACG 1706  
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGCCCTGGAGGGTCTGGAAT 1744  
Db 1707 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGGCAGATTTCTGGAGGCAATGTCTAT 1766  
QY 1745 ACCTCCACTCACGAAGGATTTGCAATGGGAGCTCAAGCTGACAAAGCTGCTCCTGTCCA 1804  
Db 1767 ACCTGCACAGCAACATGATTTGCAATGGGAGCATCAAGGAGGCCAATATCTCCAGACT 1826  
QY 1805 GCGATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864  
Db 1827 CAGCTGGGA--ATGTGAAGCTTGGGATTTTGGGCCAGCAACGCCCTACAGACCATCT 1883  
QY 1865 GCCTGGGAAGTCTTGTCTACAGGGGACTATACCTCCCTGGCAGACAGACCCACATGGCTC 1924  
Db 1884 GCATGTCCAGGACAGGCAT-----TCGCTCTGTCACTGGCACACCCCTACTGGATGATC 1937  
QY 1925 CGAGGTGTGTGGGCGAGAGCTGCGACGCCAAGGTGATGCTTGGAGCAGCTGCTGTA 1984  
Db 1938 CTGAAGTCTCATGATGGCGAGGGCTATGGAAGAGGACAGACGTGTGGAGCCTGGGCTGTA 1997  
QY 1985 TGATGCTGCACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCTTCGAGGGCGGCTCT 2044  
Db 1998 CTGTGGTGGAAATGCTGACAGAGAACCCACTTGGGCGAGATGATGAAGCTATGGCTGCTCA 2057  
QY 2045 GCCTCAAGATTGCCAGGAGCCCTCC 2069  
Db 2058 TTTTCAAGATTGCCAGCCAGCCTAC 2082

RESULT 15  
US-09-423-890-5  
; Sequence 5, Application US/09423890  
; Patent No. 6312934  
; GENERAL INFORMATION:  
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; FILE REFERENCE: CPI-085CPPC  
; CURRENT APPLICATION NUMBER: US/09/423,890  
; CURRENT FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: USN 60/078,153  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: USN 60/099,165  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1935  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (25)..(1902)  
US-09-423-890-5

Query Match  
Best Local Similarity 1.8%; Score 82.2; DB 3; Length 1935;  
Matches 233; Conservative 0; Mismatches 203; Indels 9; Gaps 2;  
QY 1625 TCAACATCTTTCAGAGCTGCTGGAAGGTGGCTCCCTGGGCGAGCTGGTCAAGGACAGG 1684  
Db 1340 TCACCATCTTTATGGAGTATATGCCAGGGGGCTCTGTAAAGACCAAGTTGAAGGCTACG 1399  
QY 1685 GCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCGAGCCCTGGAGGGTCTGGAAT 1744  
Db 1400 GAGCTCTGACAGAGAGTGTGACCCGCAAGTACACCCGGCAGATTTCTGGAGGCAATGTCTAT 1459

```
Qy 1745 ACCTCCACTCAGAGGATTCTGCATGGGACGTCAAAGCTGACAAACGTGCTCCTGTCCA 1804
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1460 ACCTGCACAGCAACATGATTGTGCATCGGACATCAAGGGAGCCAATATCCTCCGAGACT 1519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1805 GCGATGGGAGCCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG 1864
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1520 CAGCTGGGA---ATGTGAAGCTTGGGGATTTGG-----GCCCAGCAACACCTACAGA 1570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1865 GCCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCAGAGACCCACATGGCTC 1924
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1571 CCACTGCATGTCAGGACAGGCATTGCTCTGTCTACTGACACACCCCTACTGGATGATC 1630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1925 CGGAGGTGGTCTGGGCGAGGAGCTGCGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGTA 1984
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1631 CTGAAGTCATCAGTGGCGCAGGGCTATGGAAGAAAGGCAGACGTGTGGAGCCTGGGCTGTA 1690
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1985 TGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCGGCTCT 2044
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1691 CTGTGGTGGATATGCTGACAGAGAAACCACTTGGCGAGAGTATGAAGCTATGGCTGCCA 1750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2045 GCCTCAAGATTGCCAGCGAGCCTCC 2069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1751 TTTTCAAGATTGCCACCCAGCCTAC 1775
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: June 10, 2005, 20:46:52  
Job time : 511.934 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 06:13:19 ; Search time 11495.3 Seconds  
(without alignments)  
16321.674 Million cell updates/sec

Title: US-09-155-676B-6  
Perfect score: 4596  
Sequence: 1 agcggggggactgtgcctg.....gatgacaatgttaaaaaa 4596

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pna/PCRUS1\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/pna/PCRUS2\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/pna/PCRUS\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/pna/US06\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/pna/US07\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/pna/US080\_COMB.seq:\*  
7: /cgn2\_6/ptodata/1/pna/US081\_COMB.seq:\*  
8: /cgn2\_6/ptodata/1/pna/US082\_COMB.seq:\*  
9: /cgn2\_6/ptodata/1/pna/US083\_COMB.seq:\*  
10: /cgn2\_6/ptodata/1/pna/US084\_COMB.seq:\*  
11: /cgn2\_6/ptodata/1/pna/US085\_COMB.seq:\*  
12: /cgn2\_6/ptodata/1/pna/US086\_COMB.seq:\*  
13: /cgn2\_6/ptodata/1/pna/US087\_COMB.seq:\*  
14: /cgn2\_6/ptodata/1/pna/US088\_COMB.seq:\*  
15: /cgn2\_6/ptodata/1/pna/US089\_COMB.seq:\*  
16: /cgn2\_6/ptodata/1/pna/US090\_COMB.seq:\*  
17: /cgn2\_6/ptodata/1/pna/US091\_COMB.seq:\*  
18: /cgn2\_6/ptodata/1/pna/US092A\_COMB.seq:\*  
19: /cgn2\_6/ptodata/1/pna/US092B\_COMB.seq:\*  
20: /cgn2\_6/ptodata/1/pna/US093A\_COMB.seq:\*  
21: /cgn2\_6/ptodata/1/pna/US093B\_COMB.seq:\*  
22: /cgn2\_6/ptodata/1/pna/US094\_COMB.seq:\*  
23: /cgn2\_6/ptodata/1/pna/US095A\_COMB.seq:\*  
24: /cgn2\_6/ptodata/1/pna/US095B\_COMB.seq:\*  
25: /cgn2\_6/ptodata/1/pna/US095C\_COMB.seq:\*  
26: /cgn2\_6/ptodata/1/pna/US095D\_COMB.seq:\*  
27: /cgn2\_6/ptodata/1/pna/US096A\_COMB.seq:\*  
28: /cgn2\_6/ptodata/1/pna/US096B\_COMB.seq:\*  
29: /cgn2\_6/ptodata/1/pna/US096C\_COMB.seq:\*  
30: /cgn2\_6/ptodata/1/pna/US096D\_COMB.seq:\*  
31: /cgn2\_6/ptodata/1/pna/US096E\_COMB.seq:\*  
32: /cgn2\_6/ptodata/1/pna/US097A\_COMB.seq:\*  
33: /cgn2\_6/ptodata/1/pna/US097B\_COMB.seq:\*  
34: /cgn2\_6/ptodata/1/pna/US097C\_COMB.seq:\*  
35: /cgn2\_6/ptodata/1/pna/US098A\_COMB.seq:\*  
36: /cgn2\_6/ptodata/1/pna/US098B\_COMB.seq:\*  
37: /cgn2\_6/ptodata/1/pna/US098C\_COMB.seq:\*  
38: /cgn2\_6/ptodata/1/pna/US098D\_COMB.seq:\*  
39: /cgn2\_6/ptodata/1/pna/US099A\_COMB.seq:\*  
40: /cgn2\_6/ptodata/1/pna/US099B\_COMB.seq:\*  
41: /cgn2\_6/ptodata/1/pna/US099C\_COMB.seq:\*  
42: /cgn2\_6/ptodata/1/pna/US099D\_COMB.seq:\*  
43: /cgn2\_6/ptodata/1/pna/US099E\_COMB.seq:\*

44: /cgn2\_6/ptodata/1/pna/US099F\_COMB.seq:\*  
45: /cgn2\_6/ptodata/1/pna/US099G\_COMB.seq:\*  
46: /cgn2\_6/ptodata/1/pna/US100A\_COMB.seq:\*  
47: /cgn2\_6/ptodata/1/pna/US100B\_COMB.seq:\*  
48: /cgn2\_6/ptodata/1/pna/US101A\_COMB.seq:\*  
49: /cgn2\_6/ptodata/1/pna/US101B\_COMB.seq:\*  
50: /cgn2\_6/ptodata/1/pna/US102A\_COMB.seq:\*  
51: /cgn2\_6/ptodata/1/pna/US102B\_COMB.seq:\*  
52: /cgn2\_6/ptodata/1/pna/US103A\_COMB.seq:\*  
53: /cgn2\_6/ptodata/1/pna/US103B\_COMB.seq:\*  
54: /cgn2\_6/ptodata/1/pna/US104A\_COMB.seq:\*  
55: /cgn2\_6/ptodata/1/pna/US104B\_COMB.seq:\*  
56: /cgn2\_6/ptodata/1/pna/US105A\_COMB.seq:\*  
57: /cgn2\_6/ptodata/1/pna/US105B\_COMB.seq:\*  
58: /cgn2\_6/ptodata/1/pna/US106A\_COMB.seq:\*  
59: /cgn2\_6/ptodata/1/pna/US107A\_COMB.seq:\*  
60: /cgn2\_6/ptodata/1/pna/US107B\_COMB.seq:\*  
61: /cgn2\_6/ptodata/1/pna/US107C\_COMB.seq:\*  
62: /cgn2\_6/ptodata/1/pna/US107D\_COMB.seq:\*  
63: /cgn2\_6/ptodata/1/pna/US108A\_COMB.seq:\*  
64: /cgn2\_6/ptodata/1/pna/US108B\_COMB.seq:\*  
65: /cgn2\_6/ptodata/1/pna/US109A\_COMB.seq:\*  
66: /cgn2\_6/ptodata/1/pna/US109B\_COMB.seq:\*  
67: /cgn2\_6/ptodata/1/pna/US109C\_COMB.seq:\*  
68: /cgn2\_6/ptodata/1/pna/US110\_COMB.seq:\*  
69: /cgn2\_6/ptodata/1/pna/US6000\_COMB.seq:\*  
70: /cgn2\_6/ptodata/1/pna/US6001\_COMB.seq:\*  
71: /cgn2\_6/ptodata/1/pna/US6002\_COMB.seq:\*  
72: /cgn2\_6/ptodata/1/pna/US6003\_COMB.seq:\*  
73: /cgn2\_6/ptodata/1/pna/US6004\_COMB.seq:\*  
74: /cgn2\_6/ptodata/1/pna/US6005\_COMB.seq:\*  
75: /cgn2\_6/ptodata/1/pna/US6006\_COMB.seq:\*  
76: /cgn2\_6/ptodata/1/pna/US6007\_COMB.seq:\*  
77: /cgn2\_6/ptodata/1/pna/US6008\_COMB.seq:\*  
78: /cgn2\_6/ptodata/1/pna/US6009\_COMB.seq:\*  
79: /cgn2\_6/ptodata/1/pna/US6010\_COMB.seq:\*  
80: /cgn2\_6/ptodata/1/pna/US6011\_COMB.seq:\*  
81: /cgn2\_6/ptodata/1/pna/US6012\_COMB.seq:\*  
82: /cgn2\_6/ptodata/1/pna/US6013\_COMB.seq:\*  
83: /cgn2\_6/ptodata/1/pna/US6014\_COMB.seq:\*  
84: /cgn2\_6/ptodata/1/pna/US6015\_COMB.seq:\*  
85: /cgn2\_6/ptodata/1/pna/US6016\_COMB.seq:\*  
86: /cgn2\_6/ptodata/1/pna/US6017\_COMB.seq:\*  
87: /cgn2\_6/ptodata/1/pna/US6018\_COMB.seq:\*  
88: /cgn2\_6/ptodata/1/pna/US6019\_COMB.seq:\*  
89: /cgn2\_6/ptodata/1/pna/US6020\_COMB.seq:\*  
90: /cgn2\_6/ptodata/1/pna/US6021\_COMB.seq:\*  
91: /cgn2\_6/ptodata/1/pna/US6022\_COMB.seq:\*  
92: /cgn2\_6/ptodata/1/pna/US6023A\_COMB.seq:\*  
93: /cgn2\_6/ptodata/1/pna/US6023B\_COMB.seq:\*  
94: /cgn2\_6/ptodata/1/pna/US6024\_COMB.seq:\*  
95: /cgn2\_6/ptodata/1/pna/US6025\_COMB.seq:\*  
96: /cgn2\_6/ptodata/1/pna/US6026\_COMB.seq:\*  
97: /cgn2\_6/ptodata/1/pna/US6027\_COMB.seq:\*  
98: /cgn2\_6/ptodata/1/pna/US6028\_COMB.seq:\*  
99: /cgn2\_6/ptodata/1/pna/US6029\_COMB.seq:\*  
100: /cgn2\_6/ptodata/1/pna/US6030\_COMB.seq:\*  
101: /cgn2\_6/ptodata/1/pna/US6031\_COMB.seq:\*  
102: /cgn2\_6/ptodata/1/pna/US6032\_COMB.seq:\*  
103: /cgn2\_6/ptodata/1/pna/US6033\_COMB.seq:\*  
104: /cgn2\_6/ptodata/1/pna/US6034\_COMB.seq:\*  
105: /cgn2\_6/ptodata/1/pna/US6035\_COMB.seq:\*  
106: /cgn2\_6/ptodata/1/pna/US6036\_COMB.seq:\*  
107: /cgn2\_6/ptodata/1/pna/US6037\_COMB.seq:\*  
108: /cgn2\_6/ptodata/1/pna/US6038\_COMB.seq:\*  
109: /cgn2\_6/ptodata/1/pna/US6039\_COMB.seq:\*  
110: /cgn2\_6/ptodata/1/pna/US6040\_COMB.seq:\*  
111: /cgn2\_6/ptodata/1/pna/US6041\_COMB.seq:\*  
112: /cgn2\_6/ptodata/1/pna/US6042\_COMB.seq:\*  
113: /cgn2\_6/ptodata/1/pna/US6043\_COMB.seq:\*  
114: /cgn2\_6/ptodata/1/pna/US6044\_COMB.seq:\*  
115: /cgn2\_6/ptodata/1/pna/US6045\_COMB.seq:\*  
116: /cgn2\_6/ptodata/1/pna/US6046\_COMB.seq:\*

117: /cgn2\_6/ptodata/1/pna/US6047\_COMB.seq: \*  
118: /cgn2\_6/ptodata/1/pna/US6048\_COMB.seq: \*  
119: /cgn2\_6/ptodata/1/pna/US6049\_COMB.seq: \*  
120: /cgn2\_6/ptodata/1/pna/US6050\_COMB.seq: \*  
121: /cgn2\_6/ptodata/1/pna/US6051\_COMB.seq: \*  
122: /cgn2\_6/ptodata/1/pna/US6052\_COMB.seq: \*  
123: /cgn2\_6/ptodata/1/pna/US6053\_COMB.seq: \*  
124: /cgn2\_6/ptodata/1/pna/US6054\_COMB.seq: \*  
125: /cgn2\_6/ptodata/1/pna/US6055\_COMB.seq: \*  
126: /cgn2\_6/ptodata/1/pna/US6056\_COMB.seq: \*  
127: /cgn2\_6/ptodata/1/pna/US6057\_COMB.seq: \*  
128: /cgn2\_6/ptodata/1/pna/US6058\_COMB.seq: \*  
129: /cgn2\_6/ptodata/1/pna/US6059\_COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4596	100.0	4596	17	US-09-155-676-6	Sequence 6, Appli
2	4596	100.0	4596	17	US-09-155-676A-6	Sequence 6, Appli
3	4578.4	99.6	4596	1	PCT-US01-08631-15294	Sequence 15294, A
4	4578.4	99.6	4596	1	PCT-US02-18947-975	Sequence 975, App
5	4578.4	99.6	4596	2	PCT-US04-01100-975	Sequence 975, App
6	4578.4	99.6	4596	16	US-09-053-375B-1234	Sequence 1234, Ap
7	4578.4	99.6	4596	41	US-09-949-003C-437	Sequence 437, App
8	4578.4	99.6	4596	45	US-09-981-397A-17	Sequence 17, Appl
9	4578.4	99.6	4596	49	US-10-172-118-975	Sequence 975, App
10	4578.4	99.6	4596	51	US-10-283-975A-294	Sequence 294, App
11	4578.4	99.6	4596	52	US-10-342-887-975	Sequence 975, App
12	4578.4	99.6	4596	117	US-60-474-850-144	Sequence 144, App
13	4440	96.6	4584	41	US-09-949-003C-1291	Sequence 1291, Ap
14	4440	96.6	4684	47	US-10-087-192-887	Sequence 887, App
15	4440	96.6	4684	49	US-10-170-235-9152	Sequence 9152, Ap
16	4439.6	96.6	4591	97	US-60-278-258-4482	Sequence 4482, Ap
17	4438	96.6	4591	102	US-60-324-185-32844	Sequence 32844, A
18	4437.6	96.6	4684	114	US-60-443-566-226	Sequence 226, App
19	4437.6	96.6	4684	115	US-60-453-050-5384	Sequence 5384, Ap
20	4437.6	96.6	4684	115	US-60-453-135-5384	Sequence 5384, Ap
21	4437.6	96.6	4684	115	US-60-455-444-2732	Sequence 2732, Ap
22	4437.6	96.6	4684	116	US-60-465-241-2732	Sequence 2732, Ap
23	4437.6	96.6	4684	116	US-60-466-412-5384	Sequence 5384, Ap
24	4437.6	96.6	4684	118	US-60-485-450-566	Sequence 566, App
25	4437.6	96.6	4684	118	US-60-487-610-912	Sequence 912, App
26	4437.6	96.6	4684	126	US-60-563-440-566	Sequence 566, App
27	4437.6	96.6	4684	128	US-60-582-609-912	Sequence 912, App
28	4423.8	96.3	4596	86	US-60-172-373-3494	Sequence 3494, Ap
29	4245.6	92.4	4458	49	US-10-170-235-10109	Sequence 10109, A
30	4243.2	92.3	4458	114	US-60-443-566-227	Sequence 227, App
31	4243.2	92.3	4458	115	US-60-453-050-5385	Sequence 5385, Ap
32	4243.2	92.3	4458	115	US-60-453-135-5385	Sequence 5385, Ap
33	4243.2	92.3	4458	115	US-60-455-444-2733	Sequence 2733, Ap
34	4243.2	92.3	4458	116	US-60-465-241-2733	Sequence 2733, Ap
35	4243.2	92.3	4458	116	US-60-466-412-5385	Sequence 5385, Ap
36	4243.2	92.3	4458	118	US-60-485-450-567	Sequence 567, App
37	4243.2	92.3	4458	118	US-60-487-610-913	Sequence 913, App
38	4243.2	92.3	4458	126	US-60-563-440-567	Sequence 567, App
39	4243.2	92.3	4458	128	US-60-582-609-913	Sequence 913, App
40	4211.8	91.6	4487	22	US-09-496-914A-6881	Sequence 6881, Ap
41	4211.8	91.6	4487	26	US-09-560-875A-6881	Sequence 6881, Ap
42	2831.4	61.6	2842	94	US-60-248-798-532	Sequence 532, App
43	2673.4	58.2	3152	62	US-10-760-320A-4571	Sequence 4571, Ap
44	2673.4	58.2	3152	62	US-10-760-620A-4571	Sequence 4571, Ap
45	2054.6	44.7	2829	48	US-10-144-771-10006	Sequence 10006, A

ALIGNMENTS

RESULT 1  
US-09-155-676-6  
; Sequence 6, Application US/09155676  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/155.676  
; FILING DATE: 04-JAN-1999  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IL97/00117  
; FILING DATE: 01-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 117800  
; FILING DATE: 02-APR-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 119133  
; FILING DATE: 26-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25.618  
; REFERENCE/DOCKET NUMBER: WALLACH=21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4596 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-155-676-6  
Query Match 100.0%; Score 4596; DB 17; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGCGGGGGGACTGTGCCGTGTGGAACGTGTAGCTGTGAAGGTGGACTCTGTACCATTG 60  
Db 1 AGCGGGGGGACTGTGCCGTGTGGAACGTGTAGCTGTGAAGGTGGACTCTGTACCATTG 60  
QY 61 AGGATGTTTGGAGGATGATGTGTGGCAGAGGCACATATAACAGGCAGAGACCCCTTT 120  
Db 61 AGGATGTTTGGAGGATGATGTGTGGCAGAGGCACATATAACAGGCAGAGACCCCTTT 120  
QY 121 GCCCTGCTTTCTCCCCAACCCAGCTGACCTGTCTTCCCGAGGTCTCGGATTCTA 180  
Db 121 GCCCTGCTTTCTCCCCAACCCAGCTGACCTGTCTTCCCGAGGTCTCGGATTCTA 180  
QY 181 AGTGACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACAAGCCCTGGGAGATGGCAGTG 240  
Db 181 AGTGACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACAAGCCCTGGGAGATGGCAGTG 240

QY 241 ATGGAATGGCCCTGCCAGTCCCTGGCTCAGCAGTGGGCGACGAGAACTCCCC 300  
DB 241 ATGGAATGGCCCTGCCAGTCCCTGGCTCAGCAGTGGGCGACGAGAACTCCCC 300  
QY 301 AAGCAAAGGAGAGACGCGCCACTGGGGAGAAACAGAGCTCCGCTCTACAAGCTTGAG 360  
DB 301 AAGCAAAGGAGAGACGCGCCACTGGGGAGAAACAGAGCTCCGCTCTACAAGCTTGAG 360  
QY 361 GCCGTGGAGAGAGCCCTGTGTCTTCGGGAAGTGGGAGATCCTGAATGAAGTGAATACC 420  
DB 361 GCCGTGGAGAGAGCCCTGTGTCTTCGGGAAGTGGGAGATCCTGAATGAAGTGAATACC 420  
QY 421 AAGGCGACAGCCAAAGGAGGCTCCAGGCGAGGCGGAGCTGCATCTCTATCATCGCCAG 480  
DB 421 AAGGCGACAGCCAAAGGAGGCTCCAGGCGAGGCGGAGCTGCATCTCTATCATCGCCAG 480  
QY 481 GCTGAGTGTGAGAAATAGCAAGAGTTCAGGCCCACTTTTCAGAACGCAATTTTCATCGCT 540  
DB 481 GCTGAGTGTGAGAAATAGCAAGAGTTCAGGCCCACTTTTCAGAACGCAATTTTCATCGCT 540  
QY 541 GGCTCCAAACAGTACAGCCAGTCCAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCAT 600  
DB 541 GGCTCCAAACAGTACAGCCAGTCCAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCAT 600  
QY 601 GCTACAGAGGCGAAATGGCCCGTGTGTGGGAAGGAAAGCGTCGCAGCAAGCCCGG 660  
DB 601 GCTACAGAGGCGAAATGGCCCGTGTGTGGGAAGGAAAGCGTCGCAGCAAGCCCGG 660  
QY 661 AAGAAACGGAAGAAAGAGTCAAGTCCCTGGCTCATGAGGAGTGGCCTTGGCCAAA 720  
DB 661 AAGAAACGGAAGAAAGAGTCAAGTCCCTGGCTCATGAGGAGTGGCCTTGGCCAAA 720  
QY 721 CCCCTCCCGAGACCCCTGAGCAGGAGTGCACCATCCAGTGCAGGAGATGAGTCT 780  
DB 721 CCCCTCCCGAGACCCCTGAGCAGGAGTGCACCATCCAGTGCAGGAGATGAGTCT 780  
QY 781 CCATCGGGCGCCATATGTTAGAAACACCCGCGAGTTCACCAAGCCCTGGAAGAACCA 840  
DB 781 CCATCGGGCGCCATATGTTAGAAACACCCGCGAGTTCACCAAGCCCTGGAAGAACCA 840  
QY 841 GGCTTGGGCAACTCTGTTTTAAGCAGCTTGGCAGGSCCTACGCGGCTCTGCCTCGA 900  
DB 841 GGCTTGGGCAACTCTGTTTTAAGCAGCTTGGCAGGSCCTACGCGGCTCTGCCTCGA 900  
QY 901 TCAGAACTCCCAAACTGATCAGCCCTTGCAATGTCTGAAACCAAGTGTGAAACTGCAC 960  
DB 901 TCAGAACTCCCAAACTGATCAGCCCTTGCAATGTCTGAAACCAAGTGTGAAACTGCAC 960  
QY 961 CACCCCGAGGAGGCGCCCTGCGCCCTGCGCCAGCAACCCCTTCCCTATAGCAGACTG 1020  
DB 961 CACCCCGAGGAGGCGCCCTGCGCCCTGCGCCAGCAACCCCTTCCCTATAGCAGACTG 1020  
QY 1021 CCTCATCCCTTCCCATCCACCTCTCAGCCCTGGAAACCTCACCCCTCTGGAGTCTTC 1080  
DB 1021 CCTCATCCCTTCCCATCCACCTCTCAGCCCTGGAAACCTCACCCCTCTGGAGTCTTC 1080  
QY 1081 CTGGGCAAACTGGCTGTAGACAGCCAGAAACCTTGGCTTGACCCACACTGAGCAAA 1140  
DB 1081 CTGGGCAAACTGGCTGTAGACAGCCAGAAACCTTGGCTTGACCCACACTGAGCAAA 1140  
QY 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGTGCCTG 1200  
DB 1141 CTGGCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGTGCCTG 1200  
QY 1201 TCTGTGTGCGCCATGAGAACTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGC 1260  
DB 1201 TCTGTGTGCGCCATGAGAACTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGC 1260  
QY 1261 AGCGTGAAGTCAAGCCAGGCGCCACAGCTGTGACCGCTGGCAGACCTGGGCGAGCAGG 1320  
DB 1261 AGCGTGAAGTCAAGCCAGGCGCCACAGCTGTGACCGCTGGCAGACCTGGGCGAGCAGG 1320  
QY 1321 GGCTCCAGATCCCGGAGCCCGAGCCCAAAACTGAGGCAACGAGGGTGTCTCTGCTCACT 1380

DB 1321 GGCTCCAGATCCCGGAGCCCGAGCCCAAAACTGAGGCAACGAGGGTGTCTCTGCTCACT 1380  
QY 1381 GAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAGTCCACTGGGCGCACGACCAG 1440  
DB 1381 GAGAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAGTCCACTGGGCGCACGACCAG 1440  
QY 1441 CTCCGCTGGGCGAGAGGCTCTCTCGGAGAGTGCACAGGATGGAGGCAAGCAGACTGGC 1500  
DB 1441 CTCCGCTGGGCGAGAGGCTCTCTCGGAGAGTGCACAGGATGGAGGCAAGCAGACTGGC 1500  
QY 1501 TTCCAGTGGCTGTCAAAAAGTGGCTGGAGTATTTCCGGCGAGAGGAGTGTATGGCA 1560  
DB 1501 TTCCAGTGGCTGTCAAAAAGTGGCTGGAGTATTTCCGGCGAGAGGAGTGTATGGCA 1560  
QY 1561 TGTGAGGATTCACCTCACCCAGAAATTCCTCTTTGATGGAGCTGTGAGAGAAGGCTCT 1620  
DB 1561 TGTGAGGATTCACCTCACCCAGAAATTCCTCTTTGATGGAGCTGTGAGAGAAGGCTCT 1620  
QY 1621 TGGGTCAACATCTTCATGAGCTGTGGAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680  
DB 1621 TGGGTCAACATCTTCATGAGCTGTGGAAGTGGCTCCCTGGGCGAGCTGGTCAAGGAG 1680  
QY 1681 CAGGCTGTCTCCAGAGGACCGGCGCTGTACTACTTGGGCGAGGCTTGGAGGCTGTG 1740  
DB 1681 CAGGCTGTCTCCAGAGGACCGGCGCTGTACTACTTGGGCGAGGCTTGGAGGCTGTG 1740  
QY 1741 GAATACCTCCACTCAAGAGGATTCGATGGGAGCTCAAGAGCTGACAAAGTGTCTCTG 1800  
DB 1741 GAATACCTCCACTCAAGAGGATTCGATGGGAGCTCAAGAGCTGACAAAGTGTCTCTG 1800  
QY 1801 TCCAGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTCAACCT 1860  
DB 1801 TCCAGCGATGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTCAACCT 1860  
QY 1861 GATGGCTGGGAAAGTCTTGTCTCACAGGAGCTACATCTTGGCGACAGAGCCACATG 1920  
DB 1861 GATGGCTGGGAAAGTCTTGTCTCACAGGAGCTACATCTTGGCGACAGAGCCACATG 1920  
QY 1921 GCTCGGAGTGGTGTCTGGGCGAGGCTGCGACGCAAGTGGATGTCTGGAGAGCTGC 1980  
DB 1921 GCTCGGAGTGGTGTCTGGGCGAGGCTGCGACGCAAGTGGATGTCTGGAGAGCTGC 1980  
QY 1981 TGTATGATGTGCACTGCTCAACGGCTGCGACCCCTGGACTCAGTCTTCCGAGGGCG 2040  
DB 1981 TGTATGATGTGCACTGCTCAACGGCTGCGACCCCTGGACTCAGTCTTCCGAGGGCG 2040  
QY 2041 CTCTGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGGCC 2100  
DB 2041 CTCTGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGGCC 2100  
QY 2101 CCTCTCAGCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCT 2160  
DB 2101 CCTCTCAGCCAGGCGCATCAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTCT 2160  
QY 2161 GCAGCGAGCTGGGAGGGAAGTGAACCGGCGACTACAGCAAGTGGAGGCTTGAAGAGC 2220  
DB 2161 GCAGCGAGCTGGGAGGGAAGTGAACCGGCGACTACAGCAAGTGGAGGCTTGAAGAGC 2220  
QY 2221 CCTTGGAGGGGAGAAATATAAGAAACCAAGCATCCACCGCAATTAAGCCAAATTAACAC 2280  
DB 2221 CCTTGGAGGGGAGAAATATAAGAAACCAAGCATCCACCGCAATTAAGCCAAATTAACAC 2280  
QY 2281 CAGACCTTCCATCCAGCGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCGGCCAGCT 2340  
DB 2281 CAGACCTTCCATCCAGCGAGAGCTTTCCGCAAGGGCCCCAGGGCCCCGGCCAGCT 2340  
QY 2341 GAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGGCTCTCTCCCAACAGAGCCCCCAGAG 2400  
DB 2341 GAGGAGCAACAGGCGAGAGCCCTTAAGCTCCAGGCTCTCTCTCCCAACAGAGCCCCCAGAG 2400  
QY 2401 CCAAAAGTCTCTCTCTGACTTTTGAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTA 2460





## RESULT 2

US-09-155-676A-6  
; Sequence 6, Application US/09155676A  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED  
; TITLE OF INVENTION: FACTOR (TRAF), THEIR PREPARATION AND USE  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMAR, P.L.L.C.  
; STREET: 624 Ninth Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/155,676A  
; FILING DATE: 04-JAN-1999  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IL97/00117  
; FILING DATE: 01-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: IL 117800  
; FILING DATE: 02-APR-1996  
; APPLICATION NUMBER: IL 119133  
; FILING DATE: 26-AUG-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.  
; REGISTRATION NUMBER: 25,618  
; REFERENCE/DOCKET NUMBER: WALLACH-21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4596 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
US-09-155-676A-6

Query Match - 100.0%; Score 4596; DB 17; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AGCGGGGGGACTGTGCGGTGGAAACGTGTAGCTGTTGAAGGTGGACTCTGTTACCATTTG 60  
Db 1 AGCGGGGGGACTGTGCGGTGGAAACGTGTAGCTGTTGAAGGTGGACTCTGTTACCATTTG 60  
Qy 61 AGGATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGGACCCCTTT 120  
Db 61 AGGATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAGCAGGACCCCTTT 120  
Qy 121 GCCCTGCTTTTCCCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTA 180  
Db 121 GCCCTGCTTTTCCCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTA 180  
Qy 181 AGTGACCTGCTCTGTGTTTGTGCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTG 240  
Db 181 AGTGACCTGCTCTGTGTTTGTGCTCTCTCAGGATGAGCACAAAGCTGGGAGATGGCAGTG 240

Qy 241 ATGGAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGAGAAATCCCC 300  
Db 241 ATGGAATGGCTGCCAGGTGCCCTGGCTCAGCAGTGGGGCAGAGAAATCCCC 300  
Qy 301 AAGCCAAAGGAGAGACGCCGCTGAGGAGAGAAACAGAGCTCGTCTCAAGCTTGA 360  
Db 301 AAGCCAAAGGAGAGACGCCGCTGAGGAGAGAAACAGAGCTCGTCTCAAGCTTGA 360  
Qy 361 GCCGTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAGCTGATACC 420  
Db 361 GCCGTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAGCTGATACC 420  
Qy 421 AAGGGCAGCAGCAAGGAAGGCTCCGAGGCGAGGGCCAGCTGCAATCTCTATCATCGCCAG 480  
Db 421 AAGGGCAGCAGCAAGGAAGGCTCCGAGGCGAGGGCCAGCTGCAATCTCTATCATCGCCAG 480  
Qy 481 GCTGAGTGTGAGATAGCCAAAGAGTTCAGCCCACTTTTTCAGAAAGCATTTTCATCGCT 540  
Db 481 GCTGAGTGTGAGATAGCCAAAGAGTTCAGCCCACTTTTTCAGAAAGCATTTTCATCGCT 540  
Qy 541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCAT 600  
Db 541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCAT 600  
Qy 601 GCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTCCGAGCAAAAGCCCGG 660  
Db 601 GCTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAGGAAAGCGTCCGAGCAAAAGCCCGG 660  
Qy 661 AAGAAACGGAAGAAAGAGAGCTCAAAAGTCCCTGCTCATGAGGAGTGGCTTTGGCCAAA 720  
Db 661 AAGAAACGGAAGAAAGAGAGCTCAAAAGTCCCTGCTCATGAGGAGTGGCTTTGGCCAAA 720  
Qy 721 CCCCTCCAGAGCCCTGAGCAGGAGTGTGACCATCCCAAGTCCGAGGAGATGAGTCT 780  
Db 721 CCCCTCCAGAGCCCTGAGCAGGAGTGTGACCATCCCAAGTCCGAGGAGATGAGTCT 780  
Qy 781 CCCTCGGCGCCCATATGTTAGAAACAACCCCGCAGTTTCAACAAGCTCTCAAGGAACA 840  
Db 781 CCCTCGGCGCCCATATGTTAGAAACAACCCCGCAGTTTCAACAAGCTCTCAAGGAACA 840  
Qy 841 GGCCTTGGGCAACTCTGTTTAAAGCAGTGGGAGGCGCTACGCGCGCTCTGCCTGA 900  
Db 841 GGCCTTGGGCAACTCTGTTTAAAGCAGTGGGAGGCGCTACGCGCGCTCTGCCTGA 900  
Qy 901 TCAGAACTCCACAACTGATCAGCCCTTGAATGTCTGAACCAAGTGTGAAACTGCAC 960  
Db 901 TCAGAACTCCACAACTGATCAGCCCTTGAATGTCTGAACCAAGTGTGAAACTGCAC 960  
Qy 961 CACCCCGAGGAGGAGGCGCCCTGCGCCCTGCGCCAGCACCCCTTCCCTATAGCAGACTG 1020  
Db 961 CACCCCGAGGAGGAGGCGCCCTGCGCCCTGCGCCAGCACCCCTTCCCTATAGCAGACTG 1020  
Qy 1021 CCTCATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAAACCTCACCTCTGGAGTCTTTC 1080  
Db 1021 CCTCATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAAACCTCACCTCTGGAGTCTTTC 1080  
Qy 1081 CTGGGCAAACTGGGCTGTGTAGACAGCAGAAACCTTGTGCTGACCCACACTGAGCAAA 1140  
Db 1081 CTGGGCAAACTGGGCTGTGTAGACAGCAGAAACCTTGTGCTGACCCACACTGAGCAAA 1140  
Qy 1141 CTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCAGCTGCCTG 1200  
Db 1141 CTGGCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCAGCTGCCTG 1200  
Qy 1201 TCTCGTGGTGGCCATGAGAAAGTCTGTGTGAGGAAATACCTAGTGCAATGCTCTCAAGGC 1260  
Db 1201 TCTCGTGGTGGCCATGAGAAAGTCTGTGTGAGGAAATACCTAGTGCAATGCTCTCAAGGC 1260  
Qy 1261 AGCGTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTGGCCAAAGACTGGGAGCAGCG 1320  
Db 1261 AGCGTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTGGCCAAAGACTGGGAGCAGCG 1320

QY 1321 GGCTCCAGATCCCGGAGCCAGCCCCAAAACCTGAGGACAAAGAGGGTGTCTGCTCACT 1380  
 Db |||||  
 1321 GGCTCCAGATCCCGGAGCCAGCCCCAAAACCTGAGGACAAAGAGGGTGTCTGCTCACT 1380  
 QY 1381 GAGAAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAGTCCACTGGGCCACGCAACCAG 1440  
 Db |||||  
 1381 GAGAAACTCAAGCCAGTGGATTATGAGTACCGAGAAGAGTCCACTGGGCCACGCAACCAG 1440  
 QY 1441 CTCCGCTGGGAGAGGGCTCTTCGGAGAGGTGCACAGATGGAGGACAAAGCAGACTGGC 1500  
 Db |||||  
 1441 CTCCGCTGGGAGAGGGCTCTTCGGAGAGGTGCACAGATGGAGGACAAAGCAGACTGGC 1500  
 QY 1501 TTCCAGTGGCTGTCAAAAGGTGGCCCTGGGAAGTATTTTCGGGCAGAGGAGCTGATGGCA 1560  
 Db |||||  
 1501 TTCCAGTGGCTGTCAAAAGGTGGCCCTGGGAAGTATTTTCGGGCAGAGGAGCTGATGGCA 1560  
 QY 1561 TGTGAGAGATTGACCTCACCCAGAAATTGTCCCTTTGTATGGAGCTGTGAGAGAAGGGCT 1620  
 Db |||||  
 1561 TGTGAGAGATTGACCTCACCCAGAAATTGTCCCTTTGTATGGAGCTGTGAGAGAAGGGCT 1620  
 QY 1621 TGGGTCAACATCTTCATGAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGCTCAAGGAG 1680  
 Db |||||  
 1621 TGGGTCAACATCTTCATGAGAGCTGTGGAAGGTGGCTCCCTGGGCCAGCTGCTCAAGGAG 1680  
 QY 1681 CAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTG 1740  
 Db |||||  
 1681 CAGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTCTGGGCCAGGCCCTGGAGGGTCTG 1740  
 QY 1741 GAATACCTTCACTACGAAAGGATTCGCATGGGGACGTCMAAGCTGACAAACGTCCTCTG 1800  
 Db |||||  
 1741 GAATACCTTCACTACGAAAGGATTCGCATGGGGACGTCMAAGCTGACAAACGTCCTCTG 1800  
 QY 1801 TCCAGCGATGGAGCCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAAGCT 1860  
 Db |||||  
 1801 TCCAGCGATGGAGCCAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAAGCT 1860  
 QY 1861 GATGGCTGGGAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAGACCCACATG 1920  
 Db |||||  
 1861 GATGGCTGGGAAGTCTTGTCTCAGGGGACTACATCCCTGGGCACAGAGACCCACATG 1920  
 QY 1921 GCTCCGAGGTGGTGGGCAAGAGTGCAGCGCAAGAGTGGATGTCTGGAGCAGCTGC 1980  
 Db |||||  
 1921 GCTCCGAGGTGGTGGGCAAGAGTGCAGCGCAAGAGTGGATGTCTGGAGCAGCTGC 1980  
 QY 1981 TGTATGATGCTGCATGCTCAACGGCTGCACCCCTGAGCTCAGTCTTCCGAGGGCG 2040  
 Db |||||  
 1981 TGTATGATGCTGCATGCTCAACGGCTGCACCCCTGAGCTCAGTCTTCCGAGGGCG 2040  
 QY 2041 CTCTGCCCTCAAGATTGCCAGCAGGCTCCGCTGTGAGGGAGATCCACCTCTCTGCGCC 2100  
 Db |||||  
 2041 CTCTGCCCTCAAGATTGCCAGCAGGCTCCGCTGTGAGGGAGATCCACCTCTCTGCGCC 2100  
 QY 2101 CCTCTCAGCCAGGCCATCCAGAGGGGCTGAGGAAAGGCCATCCACCGCTGTCT 2160  
 Db |||||  
 2101 CCTCTCAGCCAGGCCATCCAGAGGGGCTGAGGAAAGGCCATCCACCGCTGTCT 2160  
 QY 2161 GCAGCGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGC 2220  
 Db |||||  
 2161 GCAGCGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGAGGTCTGAAGAGC 2220  
 QY 2221 CTTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCCAAAATCAAGCCAAATTACCAC 2280  
 Db |||||  
 2221 CTTTGGAGGGAGATATAAAGAACCAAGACATCCACCGCCAAAATCAAGCCAAATTACCAC 2280  
 QY 2281 CAGACCTTCATGCCAGCCGAGAGCTTTTCGCAAGAGGCCCCAGGGGCCCGGCGAGCT 2340  
 Db |||||  
 2281 CAGACCTTCATGCCAGCCGAGAGCTTTTCGCAAGAGGCCCCAGGGGCCCGGCGAGCT 2340  
 QY 2341 GAGGAGACAAAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACCCAGAGCCCCCAGAG 2400  
 Db |||||  
 2341 GAGGAGACAAAGGAGAGCCCTTAAGCTCCAGCTCTCTCTCCACCCAGAGCCCCCAGAG 2400  
 QY 2401 CCAAAACAAGTCTCTCTCTTGAATTTGAGCAAGAGGAGTCTGGGATGTGGGAACCCCTTA 2460

Db |||||  
 2401 CCAAAACAAGTCTCTCTCTTGTGACCTTGTGACAAAGAGGAGTCTGGGATGTGGGAACCCCTTA 2460  
 QY 2461 CCTGTGCTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGGAAGCA 2520  
 Db |||||  
 2461 CCTGTGCTCTCTCTGGAGCAGCCCTGCGAGAAACCCAGCTCACCAGAGCGGGAAGCA 2520  
 QY 2521 ACCGTCCCGGAGCAGGAACTGAGCAGCTGCAAAATAGAAATATTCTCTCAACAGACCTGTCC 2580  
 Db |||||  
 2521 ACCGTCCCGGAGCAGGAACTGAGCAGCTGCAAAATAGAAATATTCTCTCAACAGACCTGTCC 2580  
 QY 2581 CAGCCATTTCTCTGGAGGAGCAGGCAAAATCTCTCGTGCCTCAGCATCGACAGCCTC 2640  
 Db |||||  
 2581 CAGCCATTTCTCTGGAGGAGCAGGCAAAATCTCTCGTGCCTCAGCATCGACAGCCTC 2640  
 QY 2641 TCCCTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCCGGGACACC 2700  
 Db |||||  
 2641 TCCCTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCCGGGACACC 2700  
 QY 2701 CTGAGCTCAGCGCTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTGGAAAC 2760  
 Db |||||  
 2701 CTGAGCTCAGCGGCTACACTCTCTGGAGCAGCAGGCGGAGGCTCGAAGCTCCAGCTGGAAAC 2760  
 QY 2761 ATGCTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCC 2820  
 Db |||||  
 2761 ATGCTGTGGCCCGGGGGCGCCACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCC 2820  
 QY 2821 CAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGGAGTTCACCGGGTCAAGTCC 2880  
 Db |||||  
 2821 CAATACAGTCTCTTAATGGTGAACACCTGCAATCCGGGAGTTCACCGGGTCAAGTCC 2880  
 QY 2881 GGAGACATCGCACCTGGCATCAGCAGCAGATCCAGCTGAGGCTTCCAGCTTGTGTCACC 2940  
 Db |||||  
 2881 GGAGACATCGCACCTGGCATCAGCAGCAGATCCAGCTGAGGCTTCCAGCTTGTGTCACC 2940  
 QY 2941 AAAGACGGGACGCTTTCGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGCGAG 3000  
 Db |||||  
 2941 AAAGACGGGACGCTTTCGCTACGACATGAGAGTGCAGACTCGGGCATCGACCTGCGAG 3000  
 QY 3001 TGCACTATGGCCCTGATGGCAGCTTCGCTGAGGCTGGAGGTTCAAGCTGCGCAGCTG 3060  
 Db |||||  
 3001 TGCACTATGGCCCTGATGGCAGCTTCGCTGAGGCTGGAGGTTCAAGCTGCGCAGCTG 3060  
 QY 3061 GAGAAACGGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGACGCTTC 3120  
 Db |||||  
 3061 GAGAAACGGCCCTAACCTGCGCTCCACCGCGGCTCCACACTGCGGAAAGACGCTTC 3120  
 QY 3121 CTGCTCGGTGCAAGATGCTGCTGAAACACAGGCTCAGCGGTTCCAGGGGATTTGCCA 3180  
 Db |||||  
 3121 CTGCTCGGTGCAAGATGCTGCTGAAACACAGGCTCAGCGGTTCCAGGGGATTTGCCA 3180  
 QY 3181 GCGCCCGGCTCAGAGTGGGAAACAGGCGCTCGAGCAGCAGGTTGGGGGCAAGCAGAT 3240  
 Db |||||  
 3181 GCGCCCGGCTCAGAGTGGGAAACAGGCGCTCGAGCAGCAGGTTGGGGGCAAGCAGAT 3240  
 QY 3241 GCCTCCAGGATTTTCAACCTGAGCCCTGCGCCACCTCTGCTGAAAAAAACATCCGCCAGCT 3300  
 Db |||||  
 3241 GCCTCCAGGATTTTCAACCTGAGCCCTGCGCCACCTCTGCTGAAAAAAACATCCGCCAGCT 3300  
 QY 3301 GAAGAGACAGAGGAGGATGGCAGGATTAAGCTGGGGAAACAAAACAGGGATCTTTTCT 3360  
 Db |||||  
 3301 GAAGAGACAGAGGAGGATGGCAGGATTAAGCTGGGGAAACAAAACAGGGATCTTTTCT 3360  
 QY 3361 GCGCCCTGCTCAGTCCAGTTCGCTGACCGCTGGATTCAGTTCAGTTCAGTTCAGTTCAG 3420  
 Db |||||  
 3361 GCGCCCTGCTCAGTCCAGTTCGCTGACCGCTGGATTCAGTTCAGTTCAGTTCAGTTCAG 3420  
 QY 3421 AGGGAGAGCAGCTTCCAGCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACC 3480  
 Db |||||  
 3421 AGGGAGAGCAGCTTCCAGCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGCCCTCACC 3480  
 QY 3481 CTCCAGGCTGTGTGAGAGTGTCAAGTGTGTAAGGCGCCAAACTCAGGTTTCAGTGCAGAA 3540  
 Db |||||

Db 3481 CTCCAGGCTGCTGAGAGTGTCAAGTGTAAAGGGCCCAAACTCAGGTTAGTGCAGAA 3540  
QY 3541 CCAGGTACAGAGTATGCCCCCGGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTGGC 3600  
Db 3541 CCAGGTACAGAGTATGCCCCCGGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTGGC 3600  
QY 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATCCCTGACCTGGGAAG 3660  
Db 3601 CTCACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATCCCTGACCTGGGAAG 3660  
QY 3661 GCTCCCTGGTGAATACACCACTCTTTCAGGTTGTGCAACACACAGGTCCTGAGTTGACC 3720  
Db 3661 GCTCCCTGGTGAATACACCACTCTTTCAGGTTGTGCAACACACAGGTCCTGAGTTGACC 3720  
QY 3721 TCTGTTTACGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGTCTCAGTCCCCCAGACA 3780  
Db 3721 TCTGTTTACGCAAGGACCAAGAGGTGTGTAAAGTGAAGTGTCTCAGTCCCCCAGACA 3780  
QY 3781 TGTGCCCCCTTTGCTGCTGGCTACCACTCTTCCCAAGAGCAGGCGCCGAGCCCTTCA 3840  
Db 3781 TGTGCCCCCTTTGCTGCTGGCTACCACTCTTCCCAAGAGCAGGCGCCGAGCCCTTCA 3840  
QY 3841 GCGCCAGCAGTGCCTCCAGACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGGTGAAGGT 3900  
Db 3841 GCGCCAGCAGTGCCTCCAGACTCGCTGGCACTCAGTTCCTCTCATCTGTAAGGTGAAGGT 3900  
QY 3901 GATGAGGATATGCTGACAGAAACAGTCTGTGGATGACATGATCAGTGTCAAGGAAAG 3960  
Db 3901 GATGAGGATATGCTGACAGAAACAGTCTGTGGATGACATGATCAGTGTCAAGGAAAG 3960  
QY 3961 CAGCAGAGAGAGCTGCGGGGCCCGCCAGCCCACTATCAGTGTCCAGGTCGTGGTTCCC 4020  
Db 3961 CAGCAGAGAGAGCTGCGGGGCCCGCCAGCCCACTATCAGTGTCCAGGTCGTGGTTCCC 4020  
QY 4021 CAGCAGAGAGTGCATCAGTGTGACACTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4080  
Db 4021 CAGCAGAGAGTGCATCAGTGTGACACTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 4080  
QY 4081 TGCCGAGGCACTTTGGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCTGCTGCTG 4140  
Db 4081 TGCCGAGGCACTTTGGCACTCTGATGACCTCAAAAGCACTTTTCATGGCTGCTGCTGCTG 4140  
QY 4141 GGCAGGAGGAGGAGTGCATCTGATGAGCAGTATGCAAGCCAGGAGATGGGGTGAAGGAC 4200  
Db 4141 GGCAGGAGGAGGAGTGCATCTGATGAGCAGTATGCAAGCCAGGAGATGGGGTGAAGGAC 4200  
QY 4201 ACAGTCTTTGAGTGTCCACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4260  
Db 4201 ACAGTCTTTGAGTGTCCACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4260  
QY 4261 ATAGCAGCCCTTCCCAATGCCCCAGCTTAGCCTCTTTCAGGGGAGGAGTACTCAGGA 4320  
Db 4261 ATAGCAGCCCTTCCCAATGCCCCAGCTTAGCCTCTTTCAGGGGAGGAGTACTCAGGA 4320  
QY 4321 CTCAGTATGATTAATCAGTGTGAATTCGTGAGGGGTGTCTGAGGCTTCAACCTCTCT 4380  
Db 4321 CTCAGTATGATTAATCAGTGTGAATTCGTGAGGGGTGTCTGAGGCTTCAACCTCTCT 4380  
QY 4381 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4440  
Db 4381 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4440  
QY 4441 CTTTGTTCAGCTGTCCCAATGCCCCAGCTTAGCCTCTTTCAGGGGAGGAGGAGGAGGAGGAG 4500  
Db 4441 CTTTGTTCAGCTGTCCCAATGCCCCAGCTTAGCCTCTTTCAGGGGAGGAGGAGGAGGAGGAG 4500  
QY 4501 CACATAAGAACTGAGCTTTGGTATCTGAGAGTCTGGGTTGTAGAGAACTCTTTTGAAG 4560  
Db 4501 CACATAAGAACTGAGCTTTGGTATCTGAGAGTCTGGGTTGTAGAGAACTCTTTTGAAG 4560  
QY 4561 CAATAAGTTTGGGTTGATGACAAATGTTAAAAAA 4596  
Db 4561 CAATAAGTTTGGGTTGATGACAAATGTTAAAAAA 4596

## RESULT 3

PCT-US01-08631-15294  
; Sequence 15294, Application PC/TUS0108631  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-049  
; CURRENT APPLICATION NUMBER: PCT/US01/08631  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 15294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1117)..(1581)  
; OTHER INFORMATION: 100% homologous to Homo sapiens sperm protein 10, accession  
; OTHER INFORMATION: number M82968, Smith-Waterman Score=809.  
PCT-US01-08631-15294

Query Match 99.6%; Score 4578.4; DB 1; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 5 GGGGAGCTGTCCCGTGTGGAAACGTGTAGCTCTTGA-AGGTGGACTCTCTTACCATTGAGG 63  
Db 5 GGGGAGCTGTCCCGTGTGGAAACGTGTAGCTCTTGAAGAGTGGACTCTCTTACCATTGAGG 64  
QY 64 ATGTTTGGAGATCAGTATGTGTGGAGAGCAACATAAAACAGCAGAGACCTTTTGGC 123  
Db 65 ATGTTTGGAGATCAGTATGTGTGGAGAGCAACATAAAACAGCAGAGACCTTTTGGC 124  
QY 124 CCTGCCCTTCTCCCGCAACCCAGGCTGACCTGTGTTCTCCAGGTCCTGGATTTAAGT 183  
Db 125 CCTGCCCTTCTCCCGCAACCCAGGCTGACCTGTGTTCTCCAGGTCCTGGATTTAAGT 184  
QY 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCTTGGAGATGGCAGTGATG 243  
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCTTGGAGATGGCAGTGATG 244  
QY 244 GAAATGGCTGCCCGAGGTCCTTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCCAAG 303  
Db 245 GAAATGGCTGCCCGAGGTCCTTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCCAAG 304  
QY 304 CCAAGGAGAGAGCGCCGCACTGGGAGAGAAACAGAGCTCCGCTACAGCTTGGAGGCC 363  
Db 305 CCAAGGAGAGAGCGCCGCACTGGGAGAGAAACAGAGCTCCGCTACAGCTTGGAGGCC 364  
QY 364 GTGAGAGAGAGCGCTGTGTTCTGCGGAAAGTGGGAGATCCCTGAATGAGTGTATACCAAG 423  
Db 365 GTGAGAGAGAGCGCTGTGTTCTGCGGAAAGTGGGAGATCCCTGAATGAGTGTATACCAAG 424  
QY 424 GGCAACAGCAAGAGGCTCCGAGGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT 483  
Db 425 GGCAACAGCAAGAGGCTCCGAGGAGGCGAGCTGCCATCTCTATCATCGCCAGGCT 484  
QY 484 GAGTGTAGATAGCAAGAGTTCAGCCCACTTTTTCAGAACGCAATTTTCATCGCTGGG 543  
Db 485 GAGTGTAGATAGCAAGAGTTCAGCCCACTTTTTCAGAACGCAATTTTCATCGCTGGG 544  
QY 544 TCCAAACAGTACAGCAGTCCGAGAGTTCAGATCCCAACAAATGTGGCCCATGCT 603  
Db 545 TCCAAACAGTACAGCAGTCCGAGAGTTCAGATCCCAACAAATGTGGCCCATGCT 604  
QY 604 ACAGAGGGCAAAATGGCCCGTGTGTGTGGAGGAGAAAGCGTCGCAGCAAGAGCCCGGAAG 663

Db	605	ACAGAGGCAAAATGGCCCGTGTGTGTTGGAGGGAAGCGTCGCAGCAAAAGCCCGAAG	664	Db	1685	GGCTGTCTCCAGAGGACCGGGCCCTGTATCTACTGTGGCCAGAGCCCTGGAGGGGTCTGAA	1744
Qy	664	AAACGGAGAGAGAGAGCTCAAGTCCCTGGCTCATGACAGAGTGGCCCTTGGCCAAACCC	723	Qy	1744	TACCTCCACTCAGCAAGGATTCTCATGGGACGCTCAAAAGCTGACAAAGCTGCTCTCTGTCC	1803
Db	665	AAACGGAGAGAGAGAGCTCAAGTCCCTGGCTCATGACAGAGTGGCCCTTGGCCAAACCC	724	Db	1745	TACCTCCACTCAGCAAGGATTCTCATGGGACGCTCAAAAGCTGACAAAGCTGCTCTGTCC	1804
Qy	724	CTCCCCAGACCCCTGAGCAGAGAGCTGCACATCCCAAGTCCAGTGCAGAGAGTGAAGTCTCCA	783	Qy	1804	AGCGATGGAGCCACGACGCCCTCTGTGACTTTTGGCCATGTGTGTGTCTTCAACCTGAT	1863
Db	725	CTCCCCAGACCCCTGAGCAGAGAGCTGCACATCCCAAGTCCAGTGCAGAGAGTGAAGTCTCCA	784	Db	1805	AGCGATGGAGCCACGACGCCCTCTGTGACTTTTGGCCATGTGTGTGTCTTCAACCTGAT	1864
Qy	784	CTCGCGCCGCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTTGAAGAAACAGGC	843	Qy	1864	GGCTTGGAAAGTCTTGTCTCAGAGGGACTACATCCCTGGCAGACAGACCCACATGGCT	1923
Db	785	CTCGCGCCGCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTTGAAGAAACAGGC	844	Db	1865	GGCTTGGAAAGTCTTGTCTCAGAGGGACTACATCCCTGGCAGACAGACCCACATGGCT	1924
Qy	844	CTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGCCTCGATCA	903	Qy	1924	CCGGAGGTGGTGTGGCGCAGGAGCTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1983
Db	845	CTTGGGCAACTCTGTTTTAAGCAGCTTGGCGAGGGCCTACGGCCGGCTCTGCCTCGATCA	904	Db	1925	CCGGAGGTGGTGTGGCGCAGGAGCTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT	1984
Qy	904	GAACTCCCAAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGAAACTGCACAC	963	Qy	1984	ATGATGTGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC	2043
Db	905	GAACTCCCAAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGAAACTGCACAC	964	Db	1985	ATGATGTGTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCCGAGGGCCGCTC	2044
Qy	964	CCCCAGGACGAGGCCCTGCCCCCTGCCACGCAACCCCTTCCCTATAGCAGACTGCCT	1023	Qy	2044	TGCCTCAAGATTGCCAGGACCTCCGCTGTGAGGAGATCCACCCCTCTCGGCCCT	2103
Db	965	CCCCAGGACGAGGCCCTGCCCCCTGCCACGCAACCCCTTCCCTATAGCAGACTGCCT	1024	Db	2045	TGCCTCAAGATTGCCAGGACCTCCGCTGTGAGGAGATCCACCCCTCTCGGCCCT	2104
Qy	1024	CATCCTTCCCATTCACACCTCTCAGCCCTGGAACCTCACCCTCTGGAGTCTTCTGT	1083	Qy	2104	CTCAGAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2163
Db	1025	CATCCTTCCCATTCACACCTCTCAGCCCTGGAACCTCACCCTCTGGAGTCTTCTGT	1084	Db	2105	CTCAGAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA	2164
Qy	1084	GGCAAACTGGCCTGTGTAGACAGCAGAAACCTTGCCTGACCCACACTGAGCAAACTG	1143	Qy	2164	GCGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTTGAAGAGCCCT	2223
Db	1085	GGCAAACTGGCCTGTGTAGACAGCAGAAACCTTGCCTGACCCACACTGAGCAAACTG	1144	Db	2165	GCGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGAGGTCTTGAAGAGCCCT	2224
Qy	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGCTGCCACACACTGGAGCCAGCTGCCTGTCT	1203	Qy	2224	TGGAGGGAGAAATATAAGAACCAAGACATCCACGGCACTCAGCAAGTGGAGGTCTTGAAGAGCCCT	2283
Db	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGCTGCCACACACTGGAGCCAGCTGCCTGTCT	1204	Db	2284	ACCTTCCATGCCAGCGAGAGAGCTTTCGCCAAGGGCCCAAGGGCCCGGCGAGCTGAG	2343
Qy	1204	CGTGTGCCATGAGAAGTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC	1263	Qy	2343	ACCTTCCATGCCAGCGAGAGAGCTTTCGCCAAGGGCCCGGCGAGCTGAG	2344
Db	1205	CGTGTGCCATGAGAAGTTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC	1264	Db	2344	GAGCAACAGGCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGGCCCTTACCT	2403
Qy	1264	GTGAGCTCAAGCCAGGCCACAGCTGACAGCTGGCCCAAGACCTGGGACAGCAGGGGC	1323	Qy	2403	GAGCAACAGGCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGGCCCTTACCT	2463
Db	1265	GTGAGCTCAAGCCAGGCCACAGCTGACAGCTGGCCCAAGACCTGGGACAGCAGGGGC	1324	Db	2404	AACAGTCTCTCTTGAAGTGGAGGAGTCTGGGATGTGGGAAACCTTACCT	2464
Qy	1324	TCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGTGTCTGTCTCACTGAG	1383	Qy	2464	CTGTCTCTCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACCAGAGGGGAAAGCAACC	2523
Db	1325	TCCAGATCCCGGAGCCAGCCCAAACTGAGGACAAAGAGGTGTCTGTCTCACTGAG	1384	Db	2524	CTGTCTCTCTGGAGCCAGCCCTTCCAGAAACCCAGCTCACCAGAGGGGAAAGCAACC	2583
Qy	1384	AACTCAAGCCAGTGGATTATGAGTACCGAGAGAGTCCATGGGCGACGACAGCTC	1443	Qy	2583	GTCCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAAAGCTGTCCAG	2643
Db	1385	AACTCAAGCCAGTGGATTATGAGTACCGAGAGAGTCCATGGGCGACGACAGCTC	1444	Db	2584	GTCCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAAATTTCTCTCAAAGCTGTCCAG	2644
Qy	1444	CGCTGGGAGAGGCTCTTCCGAGAGGTGCACAGATGGAGGACAAAGCAGACTGGCTTC	1503	Qy	2644	CTGTCTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTCAGCATCGACAGCTCTCC	2703
Db	1445	CGCTGGGAGAGGCTCTTCCGAGAGGTGCACAGATGGAGGACAAAGCAGACTGGCTTC	1504	Db	2704	CTGTCTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGCTCAGCATCGACAGCTCTCC	2763
Qy	1504	CAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTTCGGGCGAGAGAGCTGATGGCATGT	1563	Qy	2764	AGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGTCTGAAGCTTCAGCTGGAAATG	2823
Db	1505	CAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTTCGGGCGAGAGAGCTGATGGCATGT	1564	Db	2824	GTGCTGGCCCGGGGGCGGCCACCCACAGCTATTTTCAATGGTGTGAAAGTCCAA	2883
Qy	1564	GAGAGTTGACCTCACCAGAAATTTGCTCTTGTATGAGAGCTGTGAGAGAGGGCTTGG	1623	Qy	2884	GTGCTGGCCCGGGGGCGGCCACCCACAGCTATTTTCAATGGTGTGAAAGTCCAA	2943
Db	1565	GAGAGTTGACCTCACCAGAAATTTGCTCTTGTATGAGAGCTGTGAGAGAGGGCTTGG	1624	Db	2944	GTGCTGGCCCGGGGGCGGCCACCCACAGCTATTTTCAATGGTGTGAAAGTCCAA	3003
Qy	1624	GTCAACATCTTATGAGAGCTGTGGAAGTGGCTCCCTGGCCAGCTGTCAAGAGCAG	1683	Qy	3004	AGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCGGAGGTCTGAAGCTTCAGCTGGAAATG	3063
Db	1625	GTCAACATCTTATGAGAGCTGTGGAAGTGGCTCCCTGGCCAGCTGTCAAGAGCAG	1684	Db	3064	GTGCTGGCCCGGGGGCGGCCACCCACAGCTATTTTCAATGGTGTGAAAGTCCAA	3123
Qy	1684	GGCTGTCTCCAGAGACCGGGCCCTGTACTACCTGGGCCAGGCCCTTGAGGGTCTGAA	1743	Qy	3124	GTGCTGGCCCGGGGGCGGCCACCCACAGCTATTTTCAATGGTGTGAAAGTCCAA	3183

```
QY 2824 ATACAGTCTCTTAATGCTGAACACCTGACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2883
DB 2825 ATACAGTCTCTTAATGCTGAACACCTGACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2884
QY 2884 GACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGTGTACACAAA 2943
DB 2885 GACATCGGCACCTGGCATCAGCAGCCAGATCCAGCTGCAGCTTTCAGCTTGTGTACACAAA 2944
QY 2944 GACGGGACGCTGTTGCTACGACATGGAGGTGCGAGATCCGGGATCGACCTGCGAGTGC 3003
DB 2945 GACGGGACGCTGTTGCTACGACATGGAGGTGCGAGATCCGGGATCGACCTGCGAGTGC 3004
QY 3004 ACATGCGCCCTGATGGCAGCTTCCCTGAGCTGGAGGTCAAGATGGCCAGCTGGAG 3063
DB 3005 ACATGCGCCCTGATGGCAGCTTCCCTGAGCTGGAGGTCAAGATGGCCAGCTGGAG 3064
QY 3064 AACAGGCCCTTAACCCCTGCCCTCCACCGCCGCTCCACACTCCCGGAAAGCAGCTTCTCTG 3123
DB 3065 AACAGGCCCTTAACCCCTGCCCTCCACCGCCGCTCCACACTCCCGGAAAGCAGCTTCTCTG 3124
QY 3124 CTCGGTGACGATGTGCTGCTGAAACACAGGCTCAGCCGTTCCCGAGGGGATTCAGGCC 3183
DB 3125 CTCGGTGACGATGTGCTGCTGAAACACAGGCTCAGCCGTTCCCGAGGGGATTCAGGCC 3184
QY 3184 CCCGGCTCA CAGTGGGAAACAGGCTCTG CAGCAGCAAGGTGGGGCAAGCAATGCC 3243
DB 3185 CCCGGCTCA CAGTGGGAAACAGGCTCTG CAGCAGCAAGGTGGGGCAAGCAATGCC 3244
QY 3244 TCCAGGATTTTACACCTGAGCCCTGCCACCTGCTGAAACACATCCGCCAGCTGAA 3303
DB 3245 TCCAGGATTTTACACCTGAGCCCTGCCACCTGCTGAAACACATCCGCCAGCTGAA 3304
QY 3304 GAGACAGAGGAGTGGCAGGATTAACCTGGGGAACAAAACAGGGGATCTTTTCTGCC 3363
DB 3305 GAGACAGAGGAGTGGCAGGATTAACCTGGGGAACAAAACAGGGGATCTTTTCTGCC 3364
QY 3364 CCTGCTCAGTGCAGTGGCTCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423
DB 3365 CCTGCTCAGTGCAGTGGCTCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424
QY 3424 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGGCCCTTCAACCTC 3483
DB 3425 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGGCCCTTCAACCTC 3484
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGGTTGAGGAGAACCA 3543
DB 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGCCCAAACTCAGGTTGAGGAGAACCA 3544
QY 3544 GGTGAGCAGGTATGCCCGCCGTTAGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3603
DB 3545 GGTGAGCAGGTATGCCCGCCGTTAGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCTC 3604
QY 3604 ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAAGATGCTGTGACCTCGGGAAGGCT 3663
DB 3605 ACCTGGCAGCTCACCCCTTTTGGGTGAGGGGAAAAGATGCTGTGACCTCGGGAAGGCT 3664
QY 3664 CCCTGGTAGAATACACACATTTTTCAGGTTGTTGCAACACAGCTCCTGAGTTGACCTCT 3723
DB 3665 CCCTGGTAGAATACACACATTTTTCAGGTTGTTGCAACACAGCTCCTGAGTTGACCTCT 3724
QY 3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCGACACATGT 3783
DB 3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCGACACATGT 3784
QY 3784 GCGCCCTTGTGCTGGCTACCACTCTTCCCGCAGAGCAGAGCCCGAGCCCTTCAGGC 3843
DB 3785 GCGCCCTTGTGCTGGCTACCACTCTTCCCGCAGAGCAGAGCCCGAGCCCTTCAGGC 3844
QY 3844 CCAGCACTGCCCGCAGACTCGCTGSCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTAT 3903
DB 3845 CCAGCACTGCCCGCAGACTCGCTGSCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTAT 3904
```

```
QY 3904 GCAGGATATGCTGACAGGAAACAGTCTGTGGATGGACATGATCAGTCTAGGAAAGCAG 3963
DB 3905 GCAGGATATGCTGACAGGAAACAGTCTGTGGATGGACATGATCAGTCTAGGAAAGCAG 3964
QY 3964 CAGAGAGAGAGCTCCGGCGCCCGCCAGCCCACTATCAGTGTCCAGCTGCTGTTCCCCAG 4023
DB 3965 CAGAGAGAGAGCTCCGGCGCCCGCCAGCCCACTATCAGTGTCCAGGTTGCTGTTCCCCAG 4024
QY 4024 AGCACAGCTCAGCATCACTGACACTCACCTGCCCTTCCCTGCGCCCTGGCCAGAGGTAATGC 4083
DB 4025 AGCACAGCTCAGCATCACTGACACTCACCTGCCCTTCCCTGCGCCCTGGCCAGAGGTAATGC 4084
QY 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGGC 4143
DB 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGGC 4144
QY 4144 AGGCGAGGGAGTCACACTGTAGGAGCATAGCAAGCAGGAGATGGGTGAAGGACACA 4203
DB 4145 AGGCGAGGGAGTCACACTGTAGGAGCATAGCAAGCAGGAGATGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGAATA 4263
DB 4205 GTCTTGAGCTGTCCACATGATGTGACTCTCTCAAACTCTTCCAGATTTCTTAAGAATA 4264
QY 4264 GCACCCCTTCCCATTTGCCAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGGACTC 4323
DB 4265 GCACCCCTTCCCATTTGCCAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGGACTC 4324
QY 4324 ACGTAGCATTAATACACTGTGAATCTGAGGGGTGTCTGCTAGCTCAACCTCTCTGGG 4383
DB 4325 ACGTAGCATTAATACACTGTGAATCTGAGGGGTGTCTGCTAGCTCAACCTCTCTGGG 4384
QY 4384 GCAGGGACGCGGAGACTCCGTGGGAGAGCTCAATCCCACTTTGTCGAAGACAGCCTT 4443
DB 4385 GCAGGGACGCGGAGACTCCGTGGGAGAGCTCAATCCCACTTTGTCGAAGACAGCCTT 4444
QY 4444 TGTCCAGCTGTCCACATGATGATGCTGTCTCCGGGGAGAGAGCCCGCCCGCCAGCAC 4503
DB 4445 TGTCCAGCTGTCCACATGATGATGCTGTCTCCGGGGAGAGAGCCCGCCCGCCAGCAC 4504
QY 4504 ATAAAGAACTGCAAGCTTGGTACTGCAAGCTGTGGGTGTGAGAACTCTTTGTAAAGCAA 4563
DB 4505 ATAAAGAACTGCAAGCTTGGTACTGCAAGCTGTGGGTGTGAGAACTCTTTGTAAAGCAA 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595
DB 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596
```

## RESULT 4

```
PCT-US02-18947-975
; Sequence 975, Application PC/TUS0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-975
```

```
Query Match          99.6%; Score 4578.4; DB 1; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

QY 5 GGGGACTGTGCGGTGTGGAACTGTAGCTGTGA-AGGTGACTCTGTATCAATTGAGG 63  
Db 5 GGGGACTGTGCGGTGTGGAACTGTAGCTGTGTGAGAGGTGACTCTGTATCAATTGAGG 64  
QY 64 ATGTTTGGAGGATGATGATGTGTGGCAGAGGCACATAAAGCAGGACGAGACCCCTTTGCC 123  
Db 65 ATGTTTGGAGGATGATGATGTGTGGCAGAGGCACATAAAGCAGGACGAGACCCCTTTGCC 124  
QY 124 CTTGCTTTTCTCCCAACCCCAAGCTGACCTGTGTTCTCCAGGTCTGGAGTTCTAAGT 183  
Db 125 CTTGCTTTTCTCCCAACCCCAAGCTGACCTGTGTTCTCCAGGTCTGGAGTTCTAAGT 184  
QY 184 GACCTGCTCTGTGTTGTTGTTCTCTCAGATGAGCACAAAGCTGGAGATGCGAGTGATG 243  
Db 185 GACCTGCTCTGTGTTGTTGTTCTCTCAGATGAGCACAAAGCTGGAGATGCGAGTGATG 244  
QY 244 GAAATGGCTTGCCCAAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCAAG 303  
Db 245 GAAATGGCTTGCCCAAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCAAG 304  
QY 304 CCAAGGAGAGACCGCGCTCTGGGGAAGAAACAGAGCTCGGTCTCAAGCTTGAGGCC 363  
Db 305 CCAAGGAGAGACCGCGCTCTGGGGAAGAAACAGAGCTCGGTCTCAAGCTTGAGGCC 364  
QY 364 GTGGAGAAAGACCTGTGTTCTGCGGAAAGTGGAGATCCGAAATGACGTGATTAACCAAG 423  
Db 365 GTGGAGAAAGACCTGTGTTCTGCGGAAAGTGGAGATCCGAAATGACGTGATTAACCAAG 424  
QY 424 GGCACAGCCAAAGAAAGGCTCCAGAGCAGGCGCAGCTGCCATCTCTATCATCGCCAGGCT 483  
Db 425 GGCACAGCCAAAGAAAGGCTCCAGAGCAGGCGCAGCTGCCATCTCTATCATCGCCAGGCT 484  
QY 484 GAGTGTGAGATAGCCAAAGTTACGCCACCTTTTCAGAACGCATTTTCATCGCTGGG 543  
Db 485 GAGTGTGAGATAGCCAAAGTTACGCCACCTTTTCAGAACGCATTTTCATCGCTGGG 544  
QY 544 TCCAAACAGTACAGCAGTCCAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT 603  
Db 545 TCCAAACAGTACAGCAGTCCAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT 604  
QY 604 ACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG 663  
Db 605 ACAGAGGGCAAAATGGCCCGTGTGTGTTGGAAAGGAAAGCGTCGACGAAAGCCCGGAAG 664  
QY 664 AAACGGAGAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCCCTTGCCCAAAACC 723  
Db 665 AAACGGAGAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGTGGCCCTTGCCCAAAACC 724  
QY 724 CTCCCAAGGACCTTGAGCAGGAGAGCTGCACCATCCAGTGACGAGGATGAGTCTCA 783  
Db 725 CTCCCAAGGACCTTGAGCAGGAGAGCTGCACCATCCCAAGTGACGAGGATGAGTCTCA 784  
QY 784 CTCCGCGCCCAATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTGAAAGAACACAGC 843  
Db 785 CTCCGCGCCCAATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTGAAAGAACACAGC 844  
QY 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGAGCTACGCGCGGCTCTGCTCGATCA 903  
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGAGCTACGCGCGGCTCTGCTCGATCA 904  
QY 904 GAACTCCCAAACTGATCAGCCCTTGCAATGCTGAAACAGTGTGAAACCTGACACAC 963  
Db 905 GAACTCCCAAACTGATCAGCCCTTGCAATGCTGAAACAGTGTGAAACCTGACACAC 964  
QY 964 CCCCAGGAGGAGGCCCCCTGCCCTTGCCACGACCCCTTCCCTATAGCAGATGCT 1023  
Db 965 CCCCAGGAGGAGGCCCCCTGCCCTTGCCACGACCCCTTCCCTATAGCAGATGCT 1024  
QY 1024 CATCCCTTCCCAATTCAGCCCTCTCAGCCCTTGAAACCTCACCCTCTGGAGTCTTCTG 1083  
Db 1025 CATCCCTTCCCAATTCAGCCCTCTCAGCCCTTGAAACCTCACCCTCTGGAGTCTTCTG 1084

QY 1084 GGCAAAATGGGCTGTGTAGACAGCAGAAAACCTTTGCTGTGACCCACACCTGAGCAAACTG 1143  
Db 1085 GGCAAAATGGGCTGTGTAGACAGCAGAAAACCTTTGCTGTGACCCACACCTGAGCAAACTG 1144  
QY 1144 GCCTGTGTAGACAGTCCAAAAGCCCTGCTGTGGCCCAACCTGTGGAGCCAGCTGTCTGTCT 1203  
Db 1145 GCCTGTGTAGACAGTCCAAAAGCCCTGCTGTGGCCCAACCTGTGGAGCCAGCTGTCTGTCT 1204  
QY 1204 CGTGGTGGCCATGAGAAAGTCTTCTGTGGAGAAATACCTAGTGTGCTCTCAAGGAGC 1263  
Db 1205 CGTGGTGGCCATGAGAAAGTCTTCTGTGGAGAAATACCTAGTGTGCTCTCAAGGAGC 1264  
QY 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTGTGGCAAGAGCTGTGGCAGCAGCGGC 1323  
Db 1265 GTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTGTGGCAAGAGCTGTGGCAGCAGCGGC 1324  
QY 1324 TCCAGATCCCGGAGCCAGCCCCAAACTGAGGACAAAGAGGTGTCTCTCTCACTGAG 1383  
Db 1325 TCCAGATCCCGGAGCCAGCCCCAAACTGAGGACAAAGAGGTGTCTCTCTCACTGAG 1384  
QY 1384 AAACCTCAAGCCAGTGGATTATGATPACCGAAGAAAGTCCACTGGGCCACACCCAGCTC 1443  
Db 1385 AAACCTCAAGCCAGTGGATTATGATPACCGAAGAAAGTCCACTGGGCCACACCCAGCTC 1444  
QY 1444 CGCCTGGGACAGAGGCTCCTTCGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1503  
Db 1445 CGCCTGGGACAGAGGCTCCTTCGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1504  
QY 1504 CAGTGTGCTGTCAAAAAGGTGGCTGTGAAGTATTTTCGGGACAGAGGACTGATGGCATGT 1563  
Db 1505 CAGTGTGCTGTCAAAAAGGTGGCTGTGAAGTATTTTCGGGACAGAGGACTGATGGCATGT 1564  
QY 1564 CAGGATGACTCTCACCCAGAAATTTGCTCCCTTTGATGTGAGCTGTGAGAAAGGCTTGG 1623  
Db 1565 CAGGATGACTCTCACCCAGAAATTTGCTCCCTTTGATGTGAGCTGTGAGAAAGGCTTGG 1624  
QY 1624 GTCAACATCTTTCATGGAGTCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1683  
Db 1625 GTCAACATCTTTCATGGAGTCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1684  
QY 1684 GGTGTGTCTCCAGAGGACCGGGCCCTGTATCTTGGGCCAGGCTCTGGAGGCTGTGAA 1743  
Db 1685 GGTGTGTCTCCAGAGGACCGGGCCCTGTATCTTACCTGGGCCAGGCTCTGGAGGCTGTGAA 1744  
QY 1744 TACCTCCACTCAGAAAGGATTCGATGGGACGTCGAAAGCTGCAACGCTCTCTGTCC 1803  
Db 1745 TACCTCCACTCAGAAAGGATTCGATGGGACGTCGAAAGCTGCAACGCTCTCTGTCC 1804  
QY 1804 AGCGATGGAGCCACGACGACCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT 1863  
Db 1805 AGCGATGGAGCCACGACGACCTCTGTGACTTTTGGCCATGCTGTGTCTTCAACCTGAT 1864  
QY 1864 GGCCTGGAAAGTCTTGTCTCAGAGGACTACATCCCTGGCACAGAACCCACATGGCT 1923  
Db 1865 GGCCTGGAAAGTCTTGTCTCAGAGGACTACATCCCTGGCACAGAACCCACATGGCT 1924  
QY 1924 CCGGAGGTGTCTGGGACAGGAGTGCAGCCCAAGTGGATGTCTGGAGCAGCTGTGT 1983  
Db 1925 CCGGAGGTGTGTGGGACAGGAGTGCAGCCCAAGTGGATGTCTGGAGCAGCTGTGT 1984  
QY 1984 ATGATGCTGCACATGCTCAAGGGCTGCCACCCCTGAGCTCAGTCTTCCGAGGGCCGCTC 2043  
Db 1985 ATGATGCTGCACATGCTCAAGGGCTGCCACCCCTGAGCTCAGTCTTCCGAGGGCCGCTC 2044  
QY 2044 TGCCTCAAGATTCAGCAGGCTCTCGCTGTGAGGAGATCCACCTCTCTGCGGCCCT 2103  
Db 2045 TGCCTCAAGATTCAGCAGGCTCTCGCTGTGAGGAGATCCACCTCTCTGCGGCCCT 2104  
QY 2104 CTCACAGCCCAAGGCTCAAGAGGGCTGAGGAAAGAGCCCATCCACGCGCTGTCTGCA 2163  
Db 2105 CTCACAGCCCAAGGCTCAAGAGGGCTGAGGAAAGAGCCCATCCACGCGCTGTCTGCA 2164  
QY 2164 GCGGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTGTGAAGAGCCCT 2223



Db 2165 ||||| CCGGAGCTGGAGGGAAGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2224  
Qy 2224 TGGAGGGGAGGAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACCAG 2283  
Db 2225 TGGAGGGGAGGAATATAAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACCAG 2284  
Qy 2284 ACCCTCCATGCCAGCCGAGAGAGGTTTGGCAAGAGGGCCCGAGGGCCCGGCCAGCTGAG 2343  
Db 2285 ACCCTCCATGCCAGCCGAGAGAGGTTTGGCAAGAGGGCCCGAGGGCCCGGCCAGCTGAG 2344  
Qy 2344 GAGACACAGCAGAGCCCTTAAGCTCCAGCTCCTCTCCACAGAGCCCGCCAGAGCCA 2403  
Db 2345 GAGACACAGCAGAGCCCTTAAGCTCCAGCTCCTCTCCACAGAGCCCGCCAGAGCCA 2404  
Qy 2404 AACAAAGTCTCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATCTGGGAACCCCTTACCT 2463  
Db 2405 AACAAAGTCTCTCCCTTGACTTTGAGCAAGAGAGAGTCTGGGATCTGGGATCTGGGATCTTACCT 2464  
Qy 2464 CTGTCTCTCCCTGGAGCCAGCCCTGCGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCTCTCCCTGGAGCCAGCCCTGCGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524  
Qy 2524 GTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATTAATTCCTCAACAGAGCTGTCCAG 2583  
Db 2525 GTCCGGAGCAGGAACCTGACAGCTGGAATAGAAATTAATTCCTCAACAGAGCTGTCCAG 2584  
Qy 2584 CCATTTTCTCTGGAGGAGCAGGCAAAATCTCTGCTGCTCAGCATCGACAGCTCTCC 2643  
Db 2585 CCATTTTCTCTGGAGGAGCAGGCAAAATCTCTGCTGCTCAGCATCGACAGCTCTCC 2644  
Qy 2644 CTGTCCGATGACAGTGAGAGAAACCCATCAAGGCTCTCAAGGCTCGCGGGAACACCTG 2703  
Db 2645 CTGTCCGATGACAGTGAGAGAAACCCATCAAGGCTCTCAAGGCTCGCGGGAACACCTG 2704  
Qy 2704 AGCTCAGGCTGACACTCTGTGAGAGCAGGAGCCAGAGCTCGAAGCTCGAGCTGGAACATG 2763  
Db 2705 AGCTCAGGCTGACACTCTGTGAGAGCAGGAGCCAGAGCTCGAAGCTCGAGCTGGAACATG 2764  
Qy 2764 GTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTTGTAAGTCCAA 2823  
Db 2765 GTGTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTTGTAAGTCCAA 2824  
Qy 2824 ATACAGTCTCTTAATGTGTGAACACTGTCACATCCGGAGTTCCACCGGGTCAAAAGTGGGA 2883  
Db 2825 ATACAGTCTCTTAATGTGTGAACACTGTCACATCCGGAGTTCCACCGGGTCAAAAGTGGGA 2884  
Qy 2884 GACATCGGCATCGGCATCAGAGCCAGATCCAGCTGAGGCTTTAGGTTGGTCAACAAA 2943  
Db 2885 GACATCGGCATCGGCATCAGAGCCAGATCCAGCTGAGGCTTTAGGTTGGTCAACAAA 2944  
Qy 2944 GACGGGAGCTGTTGCTGACGATGAGGTGCGAGCTCGGCGATCGACCTGAGTGC 3003  
Db 2945 GACGGGAGCTGTTGCTGACGATGAGGTGCGAGCTCGGCGATCGACCTGAGTGC 3004  
Qy 3004 ACATGCGCCCTGATGGCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCCAGCTGGAG 3063  
Db 3005 ACATGCGCCCTGATGGCAGCTTCGCTGAGCTGGAGGGTCAAGCATGGCCAGCTGGAG 3064  
Qy 3064 AACAGGCCCTAACCTGCTCCACCGCGGCTCCACATGCTCGGGAAGCAGCCTTCTCTG 3123  
Db 3065 AACAGGCCCTAACCTGCTCCACCGCGGCTCCACATGCTCGGGAAGCAGCCTTCTCTG 3124  
Qy 3124 CTGGGTGAGATGCTGCTGTAACACAGGCTCAGGCTTCCAGGGGATTCGACAGCC 3183  
Db 3125 CTGGGTGAGATGCTGCTGTAACACAGGCTCAGGCTTCCAGGGGATTCGACAGCC 3184  
Qy 3184 CCCGGCTCAGTGGGAACCAAGGCTCGCAGCAGCAAGAGTGGGGGCAAGCAAGTGC 3243  
Db 3185 CCCGGCTCAGTGGGAACCAAGGCTCGCAGCAGCAAGAGTGGGGGCAAGCAAGTGC 3244  
Qy 3244 TCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGTAAGAAACATCCGCCACGTGAA 3303

Db 3245 TCCAGGATTTACACCTGAGCCCTGCCCCACCTGCTGTAAAAAAACATCCGCCACGTGAA 3304  
Qy 3304 GAGACAGAGGAGGATGGCAGGATTAACCTGGGAAAAACAAACAGGGATCTTTTCTGCC 3363  
Db 3305 GAGACAGAGGAGGATGGCAGGATTAACCTGGGAAAAACAAACAGGGATCTTTTCTGCC 3364  
Qy 3364 CTTGCTCAGTCCAGTTGGCTGACCCGCTTGGATCAGTGAACCATTTGTTGGCAGACAGG 3423  
Db 3365 CTTGCTCAGTCCAGTTGGCTGACCCGCTTGGATCAGTGAACCATTTGTTGGCAGACAGG 3424  
Qy 3424 GGAGAGCAGCTTTCAGGCTTGGGTCAGAGGGGTGGCGAGCCCTTCGSCCCTCACCCCTC 3483  
Db 3425 GGAGAGCAGCTTTCAGGCTTGGGTCAGAGGGGTGGCGAGCCCTTCGSCCCTCACCCCTC 3484  
Qy 3484 CAGGCTCTGTGAGAGTGTCAAGTGTGAAGGGGCCCAAACTCAGGTTTCAGTGCAGAAACA 3543  
Db 3485 CAGGCTCTGTGAGAGTGTCAAGTGTGAAGGGGCCCAAACTCAGGTTTCAGTGCAGAAACA 3544  
Qy 3544 GGTTCAGCAGGATATGCCCGCCGCTTAAGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3603  
Db 3545 GGTTCAGCAGGATATGCCCGCCGCTTAAGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3604  
Qy 3604 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCTGACCCCTGSGGAAGGCT 3663  
Db 3605 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAGAAATGCTGACCCCTGSGGAAGGCT 3664  
Qy 3664 CCCTGGTAGAATAACACACACTTTTTCAGGTTTGTGCAACACAGGTCCTGAGTTCACCTCT 3723  
Db 3665 CCCTGGTAGAATAACACACACTTTTTCAGGTTTGTGCAACACAGGTCCTGAGTTCACCTCT 3724  
Qy 3724 GGTTCAGCCAAAGAACCAAGAGGTGTAAAGTGAAGTGTTCAGTCCCAGACATGT 3783  
Db 3725 GGTTCAGCCAAAGAACCAAGAGGTGTAAAGTGAAGTGTTCAGTCCCAGACATGT 3784  
Qy 3784 GCGCTTTGCTGTGGCTACCACTTTCCAGAGCAGCAGGCCCCAGGCCCTTCAGGC 3843  
Db 3785 GCGCTTTGCTGTGGCTACCACTTTCCAGAGCAGCAGGCCCCAGGCCCTTCAGGC 3844  
Qy 3844 CCAGCAGTCCCGCAGACTCGCTGGCACTCATCTGTAAGAGTGAAGGTTGAT 3903  
Db 3845 CCAGCAGTCCCGCAGACTCGCTGGCACTCATCTGTAAGAGTGAAGGTTGAT 3904  
Qy 3904 GCAGGATATCTGTCAGAGAACAGTCTGTGGATGGACATGATCAGTCTAAGGAAAGCAG 3963  
Db 3905 GCAGGATATCTGTCAGAGAACAGTCTGTGGATGGACATGATCAGTCTAAGGAAAGCAG 3964  
Qy 3964 CAGAGAGAGCTCCGGGCCCGCCAGCCCACTATCAGTGTCCAGGCTGCTGGTTCGCCAG 4023  
Db 3965 CAGAGAGAGCTCCGGGCCCGCCAGCCCACTATCAGTGTCCAGGCTGCTGGTTCGCCAG 4024  
Qy 4024 AGCAGCTCAGCATCACACTGACACTCACCTGCGCTGCCCTGCGCCAGAGGGTACTGC 4083  
Db 4025 AGCAGCTCAGCATCACACTGACACTCACCTGCGCTGCCCTGCGCCAGAGGGTACTGC 4084  
Qy 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGCTGCTGCGCCCTGCGCAGGGC 4143  
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGCTGCTGCGCCCTGCGCAGGGC 4144  
Qy 4144 AGGCGAGGGCAGTGAACACTGTAGAGCATAGCAAGCAGAGATGGGGTGAAGGAGACACA 4203  
Db 4145 AGGCGAGGGCAGTGAACACTGTAGAGCATAGCAAGCAGAGATGGGGTGAAGGAGACACA 4204  
Qy 4204 GTCTTGAGCTGTCACATGATGACTCTCTCAAACTCTCCAGATTTCTTAAGATA 4263  
Db 4205 GTCTTGAGCTGTCACATGATGACTCTCTCAAACTCTCTCAAGATTTCTTAAGATA 4264  
Qy 4264 GCACCCCTTTCCCATTTGCCCCAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGGACTC 4323  
Db 4265 GCACCCCTTTCCCATTTGCCCCAGCTTAGCTCTTCTCCAGGGGAGCTACTCAGGACTC 4324  
Qy 4324 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCCTCAACCTCTGGG 4383  
Db 4325 ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCCCTCAACCTCTGGG 4384





Db 1565 GCAGATTGACCTCACCAGATTGTCCTTTGTATGAGAGCTGTGAGAGAGGGGCTTGG 1624  
Qy 1624 GTCAACATCTTTCATGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGAGCAG 1683  
Db 1625 GTCAACATCTTTCATGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGAGCAG 1684  
Qy 1684 GGCTGTCTCCAGAGAGCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGGAA 1743  
Db 1685 GGCTGTCTCCAGAGAGCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGGAA 1744  
Qy 1744 TACCTCCACTCACAAGAGATTCTGATGGGACGCTCAAGAGTGAACAAGTGTCTCTCTGTCC 1803  
Db 1745 TACCTCCACTCACAAGAGATTCTGATGGGACGCTCAAGAGTGAACAAGTGTCTCTGTCC 1804  
Qy 1804 AGCGATGGGAGCCACGACGCCCTCTGTGACTTTGGCCATGTCTGTGTCTTCAACCTGAT 1863  
Db 1805 AGCGATGGGAGCCACGACGCCCTCTGTGACTTTGGCCATGTCTGTGTCTTCAACCTGAT 1864  
Qy 1864 GGCTGGGAAAGTCTTCTCTCAGAGGAGCTACATCCCTGGGCACAGAACCCATGGCT 1923  
Db 1865 GGCTGGGAAAGTCTTCTCTCAGAGGAGCTACATCCCTGGGCACAGAACCCATGGCT 1924  
Qy 1924 CCGAGGTGGTGTGGGACAGAGCTGCCAGCCGCTGCAAGGTGGATGTCTGGAGCAGTGTGT 1983  
Db 1925 CCGAGGTGGTGTGGGACAGAGCTGCCAGCCGCTGCAAGGTGGATGTCTGGAGCAGTGTGT 1984  
Qy 1984 ATGATGTGACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC 2043  
Db 1985 ATGATGTGACATGCTCAACGGGTGCCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC 2044  
Qy 2044 TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCACCTCTCTGGCCGCT 2103  
Db 2045 TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCACCTCTCTGGCCGCT 2104  
Qy 2104 CTCACGCCAGGCGCATCCAGAGGGCTGAGGAAGCCATCCACCGGTGTCTGCA 2163  
Db 2105 CTCACGCCAGGCGCATCCAGAGGGCTGAGGAAGCCATCCACCGGTGTCTGCA 2164  
Qy 2164 CGGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2223  
Db 2165 CGGAGCTGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2224  
Qy 2224 TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAAG 2283  
Db 2225 TGGAGGGGAGAAATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAAG 2284  
Qy 2284 ACCCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGCCCGGCCAGCTGAG 2343  
Db 2285 ACCCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGCCCGGCCAGCTGAG 2344  
Qy 2344 GAGACAAGCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCCA 2403  
Db 2345 GAGACAAGCAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCGCCAGAGCCCA 2404  
Qy 2404 AACAAAGTCTCTCTCTTGAATTTGAGCAAGGAGGTCTGGGATGTGGGAACCTTACCT 2463  
Db 2405 AACAAAGTCTCTCTCTTGAATTTGAGCAAGGAGGTCTGGGATGTGGGAACCTTACCT 2464  
Qy 2464 CTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCTCTCTGGAGCCAGCCCTGCCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524  
Qy 2524 GTCCCGGAGCAGGAAGTGCAGCAGCTGGAATAGAAATATTCTCAACAGCTGTCCAG 2583  
Db 2525 GTCCCGGAGCAGGAAGTGCAGCAGCTGGAATAGAAATATTCTCAACAGCTGTCCAG 2584  
Qy 2584 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCTGTGCTCTCAGCATCGACGCTCTCC 2643  
Db 2585 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCTGTGCTCTCAGCATCGACGCTCTCC 2644  
Qy 2644 CTGTCCGATGACGTGAGAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG 2703  
Db 2645 CTGTCCGATGACGTGAGAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCCCTG 2704

Qy 2704 AGCTCAGGCGTACACTCTCTGGAGCAGCCGAGGCTCGAAGCTCCAGCTCGAACAATG 2763  
Db 2705 AGCTCAGGCGTACACTCTCTGGAGCAGCCGAGGCTCGAAGCTCCAGCTCGAACAATG 2764  
Qy 2764 GTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTATTCATGTGTGAAGTCCAA 2823  
Db 2765 GTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTATTCATGTGTGAAGTCCAA 2824  
Qy 2824 ATACAGTCTCTTAATGTGAACACTGTCACATCCGGAGTTCCACCGGGTCAAGTGGGA 2883  
Db 2825 ATACAGTCTCTTAATGTGAACACTGTCACATCCGGAGTTCCACCGGGTCAAGTGGGA 2884  
Qy 2884 GACATCCGCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTGGTTCACCAAA 2943  
Db 2885 GACATCCGCACTGGCATCAGCAGCAGATCCAGCTTCAGCTTGGTTCACCAAA 2944  
Qy 2944 GACGGGAGCCTGTTGCTAGCAGATGAGAGTGCAGACTCGGGATCGAAGCTGAGTGC 3003  
Db 2945 GACGGGAGCCTGTTGCTAGCAGATGAGAGTGCAGACTCGGGATCGAAGCTGAGTGC 3004  
Qy 3004 ACCTGGCCCTGATGAGCTTCGCTGGAGCTGAGGGTCAAGCATGGCCAGCTGAG 3063  
Db 3005 ACCTGGCCCTGATGAGCTTCGCTGGAGCTGAGGGTCAAGCATGGCCAGCTGAG 3064  
Qy 3064 AACAGGCCCTAACCTGCTCCACCGCGGCTCCACACTCCCGGAAAGCAGCTTCCTG 3123  
Db 3065 AACAGGCCCTAACCTGCTCCACCGCGGCTCCACACTCCCGGAAAGCAGCTTCCTG 3124  
Qy 3124 CTGGTGCAGATGCTGCTGCTGAAAACACAGGCTCAGCGTTCCAGGGGATTCAGGCC 3183  
Db 3125 CTGGTGCAGATGCTGCTGCTGAAAACACAGGCTCAGCGTTCCAGGGATTCAGGCC 3184  
Qy 3184 CCCGGCTCAGTGGGAAACCGGGCTCCAGCAGCAAGTGGGGCAAGCAGAAAGCC 3243  
Db 3185 CCCGGCTCAGTGGGAAACCGGGCTCCAGCAGCAAGTGGGGCAAGCAGAAAGCC 3244  
Qy 3244 TCCAGGATTTTCACTCAGCTGAGCCCTGCCCAACCTGCTGAAAAAACAATCCGCCACGTGA 3303  
Db 3245 TCCAGGATTTTCACTCAGCTGAGCCCTGCCCAACCTGCTGAAAAAACAATCCGCCACGTGA 3304  
Qy 3304 GAGACAGAGAGGATGGCAGGAGTTACTGGGAAAACAAACAGGGATTTTTTCTGCC 3363  
Db 3305 GAGACAGAGGAGGATGGCAGGAGTTACTGGGAAAACAAACAGGGATTTTTTCTGCC 3364  
Qy 3364 CCTGCTCAGTTCAGTGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG 3423  
Db 3365 CCTGCTCAGTTCAGTGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG 3424  
Qy 3424 GGAGAGCAGCTTCCAGCTGGGTGAGAGGGTGGCGAGCCCTTCGGCCCTCACCCCTC 3483  
Db 3425 GGAGAGCAGCTTCCAGCTGGGTGAGAGGGTGGCGAGCCCTTCGGCCCTCACCCCTC 3484  
Qy 3484 CAGGCTCTGTGAGAGTGTCAAGTGTCTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3543  
Db 3485 CAGGCTCTGTGAGAGTGTCAAGTGTCTAAGGGGCCCTCTAAGGGGCCCTTGCCTGGCCCTC 3544  
Qy 3544 GGTTCAGAGTATGCCCGCCGCTAGGTTAAGGGGCCCTCTAAGGGGCCCTTGCCTGGCCCTC 3603  
Db 3545 GGTTCAGAGTATGCCCGCCGCTAGGTTAAGGGGCCCTCTAAGGGGCCCTTGCCTGGCCCTC 3604  
Qy 3604 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAAATGCTGACCTTGGGAAGGCT 3663  
Db 3605 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAAATGCTGACCTTGGGAAGGCT 3664  
Qy 3664 CCCTGGTAGAATACACCACTTTTTCAGGTTTGTGCAACACAGGCTCTCTGAGTTGACCTCT 3723  
Db 3665 CCCTGGTAGAATACACCACTTTTTCAGGTTTGTGCAACACAGGCTCTCTGAGTTGACCTCT 3724  
Qy 3724 GGTTCAGCAAGGACCAAGAGAGTGTGAAGTGAAGTGTCTCAGTCTCCCAAGCATGT 3783  
Db 3725 GGTTCAGCAAGGACCAAGAGAGTGTGAAGTGAAGTGTCTCAGTCTCCCAAGCATGT 3784

```
QY 3784 GCCCTTTCTGCTGCTACCACTCTTCCCAAGACAGAGAGCCCGGAGCCCTTCAGGC 3843
Db 3785 GCCCTTTGCTGCTGGCTACCACTCTTCCCAAGACAGAGAGCCCGGAGCCCTTCAGGC 3844
QY 3844 CCAGCACTGCCACAGCTCGCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTGAT 3903
Db 3845 CCAGCACTGCCACAGCTCGCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTGAT 3904
QY 3904 CGAGGATATGCTGTACAGAAAGTCTGTGGATGACATGATCAGTCTTAAGGAAAGCAG 3963
Db 3905 CGAGGATATGCTGTACAGAAAGTCTGTGGATGACATGATCAGTCTTAAGGAAAGCAG 3964
QY 3964 CAGAGAGAGAGCTCCGGGCGCCCGACCCCACTATCAGTGTCCAGGCTGTGTTCCCGAG 4023
Db 3965 CAGAGAGAGAGCTCCGGGCGCCCGACCCCACTATCAGTGTCCAGGCTGTGTTCCCGAG 4024
QY 4024 AGCAGAGTACAGATCAGTGTACACTGACACTGACCTCAAGCACTTTTCATGGCTGCCCTTGGCAGGCG 4083
Db 4025 AGCAGAGTACAGATCAGTGTACACTGACACTGACCTCAAGCACTTTTCATGGCTGCCCTTGGCAGGCG 4084
QY 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTTGGCAGGCG 4143
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTTGGCAGGCG 4144
QY 4144 AGGGCAGGGCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4203
Db 4145 AGGGCAGGGCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGTGTCCACATGATGATGATCTCTCAAGCTTTCAGATTTCTTAAGAATA 4263
Db 4205 GTCTTGAGTGTCCACATGATGATGATCTCTCAAGCTTTCAGATTTCTTAAGAATA 4264
QY 4264 GCACCCCTTCCCACTGCCCCAGCTTAGCCCTTCTCCAGGGGAGCTACTCAGGACTC 4323
Db 4265 GCACCCCTTCCCACTGCCCCAGCTTAGCCCTTCTCCAGGGGAGCTACTCAGGACTC 4324
QY 4324 AGCTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCCTGGG 4383
Db 4325 AGCTAGCATTAATCAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCCTGGG 4384
QY 4384 GAGGGGAGCCGAGACTCCGTGGGAGAGCTCATTCCACATCTTGCCAGACAGCCCTT 4443
Db 4385 GAGGGGAGCCGAGACTCCGTGGGAGAGCTCATTCCACATCTTGCCAGACAGCCCTT 4444
QY 4444 TGTCCAGCTGTCCACATGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCACATGAGTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAA 4564
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595
Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596
```

```
RESULT 6
US-09-053-375B-1234
; Sequence 1234, Application US/09053375B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: Nucleic Acid Arrays
; FILE REFERENCE: CLON-006
; CURRENT APPLICATION NUMBER: US/09/053,375B
; CURRENT FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 1543
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1234
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-053-375B-1234
Query Match
Best Local Similarity 99.6%; Score 4578.4; DB 16; Length 4596;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 5 GGGGACCTGTCCCGTGTGGAACGTGTAGCTTTGA-AGGTGACATCTGTTACCATTTGAGG 63
Db 5 GGGGACCTGTCCCGTGTGGAACGTGTAGCTTTGAGAGGTGACATCTCTTACCATTTGAGG 64
QY 64 ATGTTTGGAGATGATGATGTGTGGCAGAGGACACATATAACAGCAGAGACCCCTTTGCC 123
Db 65 ATGTTTGGAGATGATGATGTGTGGCAGAGGACACATATAACAGCAGAGACCCCTTTGCC 124
QY 124 CCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCAGAGTCTGGGATTTCTAAGT 183
Db 125 CCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCAGAGTCTGGGATTTCTAAGT 184
QY 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCTTGGGAGATGGCAGTATG 243
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCACAAGCTTGGGAGATGGCAGTATG 244
QY 244 GAAATGGGCTGCCCAAGTGGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCAAG 303
Db 245 GAAATGGGCTGCCCAAGTGGCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCAAG 304
QY 304 CCAAGAGAGAAAGACGCGCCCACTGGGGAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 363
Db 305 CCAAGAGAGAAAGACGCGCCCACTGGGGAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC 364
QY 364 GTGGAGAAAGACCCCTGTGTTCTGCGGAAAGTGGGAGATCTTGAAATGAAGTAAATCCAAAG 423
Db 365 GTGGAGAAAGACCCCTGTGTTCTGCGGAAAGTGGGAGATCTTGAAATGAAGTAAATCCAAAG 424
QY 424 GGCACAGCAAGAAAGGCTCCGAGCAGGCGCAGCTGCTCTCTCATCGCCAGGCT 483
Db 425 GGCACAGCAAGAAAGGCTCCGAGCAGGCGCAGCTGCTCTCTCATCGCCAGGCT 484
QY 484 GAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTTCAGAAACGATTTTTCATCGCTGGG 543
Db 485 GAGTGTGAGAAATAGCAAGAGTTTCAAGCCCACTTTTTCAGAAACGATTTTTCATCGCTGGG 544
QY 544 TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCAACATGTGTGCCCATGCT 603
Db 545 TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCCAACATGTGTGCCCATGCT 604
QY 604 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGGAAGCGTCGAGCAAAAGCCCGGAAG 663
Db 605 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGGAAGCGTCGAGCAAAAGCCCGGAAG 664
QY 664 AAAAGGAAAGAAAGAGCTCAAAAGTCCCTGGCTCATGAGGAGTGGCCCTTGCCCAAAACC 723
Db 665 AAAAGGAAAGAAAGAGCTCAAAAGTCCCTGGCTCATGAGGAGTGGCCCTTGCCCAAAACC 724
QY 724 CTCCCAAGGACCCCTGAGCAGGAGTGCACCACTCCAGTCCAGTGCAGGAGATGAGTCTCA 783
Db 725 CTCCCAAGGACCCCTGAGCAGGAGTGCACCACTCCAGTCCAGTGCAGGAGATGAGTCTCA 784
QY 784 CTCGGCGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCCTCTGAGGAAACGAGC 843
Db 785 CTCGGCGCCCAATATGTTAGAAACACCCCGCAGTTCACCAAGCCCTCTGAGGAAACGAGC 844
QY 844 CTTGGGCAACTCTGTTTTAAGCAGCTTGGCAGGCGCTACGCGCCGGCTCTGCTCGATCA 903
Db 845 CTTGGGCAACTCTGTTTTAAGCAGCTTGGCAGGCGCTACGCGCCGGCTCTGCTCGATCA 904
QY 904 GAACTCCCAAACTGATCAGCCCTTGAATGTTGAAACCAAGTGTGGAATCTGCACCAAC 963
Db 905 GAACTCCCAAACTGATCAGCCCTTGAATGTTGAAACCAAGTGTGGAATCTGCACCAAC 964
QY 964 CCCCAGGAGGAGGCGCCCTTGCCTGCCACGACCCCTTCCCTATAGCAGAGTGCCT 1023
Db 965 CCCCAGGAGGAGGCGCCCTTGCCTGCCACGACCCCTTCCCTATAGCAGAGTGCCT 1024
```

QY 1024 CATCCCTTCCATTCCACCCCTCTCAGCCCTGGAACCCCTCAACCTCTGGAATCTTCTG 1083  
DB CATCCCTTCCATTCCACCCCTCTCAGCCCTGGAACCCCTCAACCTCTGGAATCTTCTG 1084  
QY 1084 GGCAGATCGGCTGTGTAGACAGCCAGAAAACCTTGCCTGACCCACACCTCTGAGCAACTG 1143  
DB GGCAGATCGGCTGTGTAGACAGCCAGAAAACCTTGCCTGACCCACACCTCTGAGCAACTG 1144  
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACCTGGAGCCAGCTGCTGTCT 1203  
DB GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACACCTGGAGCCAGCTGCTGTCT 1204  
QY 1204 CGTGTGCCCCATGAGAAATTTCTGTGTGAGAAATACCTAGTGCATGTCTCTCAAGGCAGC 1263  
DB CGTGTGCCCCATGAGAAATTTCTGTGTGAGAAATACCTAGTGCATGTCTCTCAAGGCAGC 1264  
QY 1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTGCGCCCAAGAGCTGGGACAGCGGGC 1323  
DB GTGAGCTCAAGCCAGGCCCCACAGCTGACCAAGCTGCGCCCAAGAGCTGGGACAGCGGGC 1324  
QY 1324 TCCAGATCCCGGAGCCAGCCCAAAACTGAGACAAAGAGGTGTCTCTCACTGAG 1383  
DB TCCAGATCCCGGAGCCAGCCCAAAACTGAGACAAAGAGGTGTCTCTCACTGAG 1384  
QY 1384 AAATCAAGCCAGTGGATTATGAGTACCGAAGAAAGTCCATGGGCCACGCAACGCTC 1443  
DB AAATCAAGCCAGTGGATTATGAGTACCGAAGAAAGTCCATGGGCCACGCAACGCTC 1444  
QY 1444 CGCTGGGACAGAGCTCTCGGAGAGGTGCACAGGATGGAGGACAGAGACATGGCTTC 1503  
DB CGCTGGGACAGAGCTCTCGGAGAGGTGCACAGGATGGAGGACAGAGACATGGCTTC 1504  
QY 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTGGGGCAGAGGAGCTGATGGCATGT 1563  
DB CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTGGGGCAGAGGAGCTGATGGCATGT 1564  
QY 1564 GCAGGATTGACCTCAACCCAGAAATTTGCCCTTTGTATGGAGCTGTGAGAAAGGCTTGG 1623  
DB GCAGGATTGACCTCAACCCAGAAATTTGCCCTTTGTATGGAGCTGTGAGAAAGGCTTGG 1624  
QY 1624 GTCAACATCTTATGGAGCTGCTGGAAGGTGGCTCTCGGACAGAGTGGTCAAGAGGAG 1683  
DB GTCAACATCTTATGGAGCTGCTGGAAGGTGGCTCTCGGACAGAGTGGTCAAGAGGAG 1684  
QY 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGSCAGAGCCCTGGAGGGTCTGAA 1743  
DB GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGSCAGAGCCCTGGAGGGTCTGAA 1744  
QY 1744 TACCTCCACTCAAGAGGATTTGATGGGACCGTCAAAAGCTGACAAAGCTGCTCTGTCC 1803  
DB TACCTCCACTCAAGAGGATTTGATGGGACCGTCAAAAGCTGACAAAGCTGCTCTGTCC 1804  
QY 1804 AGCAGTGGAGCCACGACGCCCTCTGTGACTTTGGCCCATGTGTGTCTTCAACCTGAT 1863  
DB AGCAGTGGAGCCACGACGCCCTCTGTGACTTTGGCCCATGTGTGTCTTCAACCTGAT 1864  
QY 1864 GGCTGTGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGCACAGAGCCCAATGGCT 1923  
DB GGCTGTGGAAAGTCTTGTCTCACAGGGACTACATCCCTGGCACAGAGCCCAATGGCT 1924  
QY 1924 CCGAGGTGGTGTGGGAGGAGTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT 1983  
DB CCGAGGTGGTGTGGGAGGAGTGCAGCCCAAGGTGGATGTCTGGAGCAGCTGTGT 1984  
QY 1984 ATGATGTGACATGCTCAACGGGTGACCCCTGGACTCAAGTCTTTCGAGGCCGCTC 2043  
DB ATGATGTGACATGCTCAACGGGTGACCCCTGGACTCAAGTCTTTCGAGGCCGCTC 2044  
QY 2044 TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCTCTCTGGCCCT 2103  
DB TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCTCTCTGGCCCT 2104

QY 2104 CTCAAGCCAGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163  
DB CTCAAGCCAGGCGCATCCAAAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164  
QY 2164 GCGAGCTGGAGGGAAGGTGAAACCGGSCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2223  
DB GCGAGCTGGAGGGAAGGTGAAACCGGSCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2224  
QY 2224 TGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCA 2283  
DB TGGAGGGGAGAAATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCA 2284  
QY 2284 ACCCTCATGCCAGCCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCCGCCAGCTGAG 2343  
DB ACCCTCATGCCAGCCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCCGCCAGCTGAG 2344  
QY 2344 GAGACACAGCAGAGAGCCCTTAAGCTCAGAGCTCTCTCCACAGAGGCCCCAGAGGCA 2403  
DB GAGACACAGCAGAGAGCCCTTAAGCTCAGAGCTCTCTCTCCACAGAGGCCCCAGAGGCA 2404  
QY 2404 AACAAAGTCTCTCTCTCTTGAATTTGAGCAAGAGGAGTCTGGGATGTGGGAAACCTTACCT 2463  
DB AACAAAGTCTCTCTCTCTTGAATTTGAGCAAGAGGAGTCTGGGATGTGGGAAACCTTACCT 2464  
QY 2464 CTGTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAGCAAC 2523  
DB CTGTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAAGCAAC 2524  
QY 2524 GTCCCGGAGCAGGAACTGACAGAGCTGGAATAGAAATTTCTCTCAAGAGCTCTCCAG 2583  
DB GTCCCGGAGCAGGAACTGACAGAGCTGGAATAGAAATTTCTCTCAAGAGCTCTCCAG 2584  
QY 2584 CCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGAGCTCTCC 2643  
DB CCATTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTCCTCAGCATCGAGCTCTCC 2644  
QY 2644 CTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACACCTG 2703  
DB CTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCGCGGACACCTG 2704  
QY 2704 AGCTCAGCGTACACTCTCTGGAGCAGCAGGCCGAGGCTCGAAGCTCAGCTGGAACATG 2763  
DB AGCTCAGCGTACACTCTCTGGAGCAGCAGGCCGAGGCTCGAAGCTCAGCTGGAACATG 2764  
QY 2764 GTGTGCGCCCGGGGGCGGCCACCGACACCCAGCTATTTCAATGTGTGAAAGTCCAA 2823  
DB GTGTGCGCCCGGGGGCGGCCACCGACACCCAGCTATTTCAATGTGTGAAAGTCCAA 2824  
QY 2824 ATACAGTCTCTTAATGTGAACACCTGCAATCCGGAGTTCCACCGGTTCAAGTGGGA 2883  
DB ATACAGTCTCTTAATGTGAACACCTGCAATCCGGAGTTCCACCGGTTCAAGTGGGA 2884  
QY 2884 GACATCCGCACTGGCATCAGAGCCAGATCCAGCTGAGGCTTTCAGTTGGTCAACCAA 2943  
DB GACATCCGCACTGGCATCAGAGCCAGATCCAGCTGAGGCTTTCAGTTGGTCAACCAA 2944  
QY 2944 GAGCGGAGCCTGTGTGCTACGATGAGGTGCGAGCTCGGCGATCGAGCTGCAAGTGC 3003  
DB GAGCGGAGCCTGTGTGCTACGATGAGGTGCGAGCTCGGCGATCGAGCTGCAAGTGC 3004  
QY 3004 ACATGGCCCTCTGATGGCAGCTTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGCTGAG 3063  
DB ACATGGCCCTCTGATGGCAGCTTTCGCTGGAGCTGGAGGTTCAAGCATGGCCAGCTGAG 3064  
QY 3064 AACAGGCCCTTAAACCTTCCCTTCAACCGCGGCTTCACTGTCCGGAAGAGCCTTCTCTG 3123  
DB AACAGGCCCTTAAACCTTCCCTTCAACCGCGGCTTCACTGTCCGGAAGAGCCTTCTCTG 3124  
QY 3124 CTGGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3183  
DB CTGGTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3184  
QY 3184 CCCCGGCTCAGTGGGAAACAGGGGCTCGCAGCAGCAAGTGGGGGCAAGCAGATGCC 3243

Db 3185 CCCCCGCTCAGTGGGAACAGGCGCTCGCAGCAGCAAGGTGGGGCAAGCAGATGCC 3244  
QY 3244 TCCAGGATTTACACCTTGAGCCCTGCCCCACCTTGCTGTAAGAAAACATCCCGCAGGTGAA 3303  
Db 3245 TCCAGGATTTACACCTTGAGCCCTGCCCCACCTTGCTGTAAGAAAACATCCCGCAGGTGAA 3304  
QY 3304 GAGCAGAGGAGGATGCGCAGAGTTACTGGGGAACAAACACAGGATCTTTTCTGCC 3363  
Db 3305 GAGCAGAGGAGGATGCGCAGAGTTACTGGGGAACAAACACAGGATCTTTTCTGCC 3364  
QY 3364 CCGTCTCCAGTCCAGTTCGCTGACCCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423  
Db 3365 CCGTCTCCAGTCCAGTTCGCTGACCCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424  
QY 3424 GGAGAGCAGCTTCAGCCTTGCGTTCAGAGGGGTGGCGAGCCCTTCGGCCCTTCACCCCTC 3483  
Db 3425 GGAGAGCAGCTTCAGCCTTGCGTTCAGAGGGGTGGCGAGCCCTTCGGCCCTTCACCCCTC 3484  
QY 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCAACTCAGGTTTCAGTGCAGAACCA 3543  
Db 3485 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCCAACTCAGGTTTCAGTGCAGAACCA 3544  
QY 3544 GGTTCAGCAGGTATGCCCCCGCTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTGGCCCTC 3603  
Db 3545 GGTTCAGCAGGTATGCCCCCGCTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTGGCCCTC 3604  
QY 3604 ACTGGCCAGCTACCCCTTTTGGGTAGAGGGGAAAGATGCCCTGACCCCTGGGAAGGCT 3663  
Db 3605 ACTGGCCAGCTACCCCTTTTGGGTAGAGGGGAAAGATGCCCTGACCCCTGGGAAGGCT 3664  
QY 3664 CCCTGGTAGAATACACCACTTTTTCAGGTTGCTGCAACACAGGTCCTGAGTTGACCTCT 3723  
Db 3665 CCCTGGTAGAATACACCACTTTTTCAGGTTGCTGCAACACAGGTCCTGAGTTGACCTCT 3724  
QY 3724 GGTTCAGCCAGGACCAAGAGGTGTGAAGTGAAGTGTCTCAGTCCCCAGACATGT 3783  
Db 3725 GGTTCAGCCAGGACCAAGAGGTGTGAAGTGAAGTGTCTCAGTCCCCAGACATGT 3784  
QY 3784 GCGCCCTTTCGCTGCTGACCACTCTTCCCAAGAGAGCAGCCCGAGCCCTTCAGGC 3843  
Db 3785 GCGCCCTTTCGCTGCTGACCACTCTTCCCAAGAGAGCAGCCCGAGCCCTTCAGGC 3844  
QY 3844 CAGACACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGAT 3903  
Db 3845 CAGACACTGCCCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGTGAT 3904  
QY 3904 GCAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAG 3963  
Db 3905 GCAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAG 3964  
QY 3964 CAGAGAGAGCTGCGGGGCCCGAGCCCACTATCAGTGTTCAGCGTGTGTTCCCCAG 4023  
Db 3965 CAGAGAGAGCTGCGGGGCCCGAGCCCACTATCAGTGTTCAGCGTGTGTTCCCCAG 4024  
QY 4024 AGCAGAGCTCAGATCAGACTGACACTCAGCTCCCTGCGCCCTGCGCAGAGGTAAGTGC 4083  
Db 4025 AGCAGAGCTCAGATCAGACTGACACTCAGCTCCCTGCGCCCTGCGCAGAGGTAAGTGC 4084  
QY 4084 CGACGGCACTTTGCACTCTGATGACCTCAAAGCACTTTTCATGGCTGCCCCCTCTGGCAGGGC 4143  
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAAGCACTTTTCATGGCTGCCCCCTCTGGCAGGGC 4144  
QY 4144 AGGGCAGGGCAGTGACACTGTAGGAGCATAGCAGCAGGATGGGGTGAAGGACACA 4203  
Db 4145 AGGGCAGGGCAGTGACACTGTAGGAGCATAGCAGCAGGATGGGGTGAAGGACACA 4204  
QY 4204 GTCTTCAGCTGTCCATGATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGATA 4263  
Db 4205 GTCTTCAGCTGTCCATGATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGATA 4264  
QY 4264 GCACCCCTTCCCAATTTGCCCAAGTTAGCCTCTCTCCAGGGGAGCTACTCAGGACTC 4323

Db 4265 GCACCCCTTCCCAATTTGCCCAAGTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC 4324  
QY 4324 ACGTAGCATTAATAACAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCTCAACCTCTCTGGG 4383  
Db 4325 ACGTAGCATTAATAACAGCTGTGAATCGTCAGGGGGTGTCTGCTAGCTCAACCTCTCTGGG 4384  
QY 4384 GCAGGGACGCGGAGACTCCGTGGGAGAAGCTCAATCCCAATCTTGCCTCAAGACAGCCTT 4443  
Db 4385 GCAGGGACGCGGAGACTCCGTGGGAGAAGCTCAATCCCAATCTTGCCTCAAGACAGCCTT 4444  
QY 4444 TGTCCAGCTGCCACATTCAGTTCAGTCTCCCGGGAGAGAGCCCGGCCCGCCAGCAC 4503  
Db 4445 TGTCCAGCTGCCACATTCAGTTCAGTCTCCCGGGAGAGAGCCCGGCCCGCCAGCAC 4504  
QY 4504 ATAAAGAACTGCAGCCTTGGTACTGCAAGTCTGGTGTGTAGAGAACTCTTTGTAAGCAA 4563  
Db 4505 ATAAAGAACTGCAGCCTTGGTACTGCAAGTCTGGTGTGTAGAGAACTCTTTGTAAGCAA 4564  
QY 4564 TAAAGTTTGGGGTGTATGACAAATGTTAAAAA 4595  
Db 4565 TAAAGTTTGGGGTGTATGACAAATGTTAAAAA 4596  
RESULT 7  
US-09-949-003C-437  
; Sequence 437, Application US/09949003C  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH NEUROPSYCHIATRIC  
; FILE REFERENCE: CL000791  
; CURRENT APPLICATION NUMBER: US/09/949,003C  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: 60/231,446  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 74065  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 437  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-949-003C-437  
Query Match 99.6%; Score 4578.4; DB 41; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 5 GGGGGAAGTGTGCGGTGGAAACGTGTAGCTGTGTA-AGGTGACTCTCTTACCATTGAGG 63  
Db 5 GGGGGAAGTGTGCGGTGGAAACGTGTAGCTGTGTAAGAGGTGGACTCTCTTACCATTGAGG 64  
QY 64 ATGTTTGGAGATGAGTATGTGGCAGAGGCACATATAACAGGCAGAGACCTTTTGGC 123  
Db 65 ATGTTTGGAGATGAGTATGTGGCAGAGGCACATATAACAGGCAGAGACCTTTTGGC 124  
QY 124 CCGTCTCTTCTCCCCCAACCCCAAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTAAAGT 183  
Db 125 CCGTCTCTTCTCCCCCAACCCCAAGGCTGACCTGTGTTCTCCAGGTCTGGGATTTAAAGT 184  
QY 184 GACCTGCTCTGTGTTTGGTCTCTTCAGATGAGCACAAGCCTGGAGATGGCAGTATG 243  
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACAAGCCTGGAGATGGCAGTATG 244  
QY 244 GAAATGGCTCTGCCAGGTGCCCTCGCTCAGCAGTGGGCGCAGCAGAGAACTCCCCCAAG 303  
Db 245 GAAATGGCTCTGCCAGGTGCCCTCGCTCAGCAGTGGGCGCAGCAGAGAACTCCCCCAAG 304  
QY 304 CCAAGGAGAGACGCGCCCACTGGGGAGAAACAGAGCTCCGTCTTAAAGTTTGAGGCC 363  
Db 305 CCAAGGAGAGACGCGCCCACTGGGGAGAAACAGAGCTCCGTCTTAAAGTTTGAGGCC 364  
QY 364 GTGGAGAGAGCCCTGTGTTCTGGCGAAGTGGGAGATCCTGAATGACGTGATTACCAAG 423

Db 365 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTAACCAAG 424  
Qy 424 GGCACAGCAAGAAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCCAAGCT 483  
Db 425 GGCACAGCAAGAAAGGCTCCGAGCAGGGCCAGCTGCCATCTCTATCATCGCCCAAGCT 484  
Qy 484 GAGCTGTGAGATAGCAAGAGTTCAAGCCCACTTTTTCAGAACGCAITTTTATCGCTGGG 543  
Db 485 GAGCTGTGAGATAGCAAGAGTTCAAGCCCACTTTTTCAGAACGCAITTTTATCGCTGGG 544  
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 603  
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGCCCATGCT 604  
Qy 604 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 663  
Db 605 ACAGAGGCAAAATGGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 664  
Qy 664 AAACGGAAAGAGAGCTCAAGTCCCTGGCTCATGAGAGTGGCCCTTGGCCAAACCC 723  
Db 665 AAACGGAAAGAGAGCTCAAGTCCCTGGCTCATGAGAGTGGCCCTTGGCCAAACCC 724  
Qy 724 CTCCCAAGACCCCTGAGCAGAGAGTGCACCAATCCCAAGTGCAGAGGATGAGTCTCCA 783  
Db 725 CTCCCAAGACCCCTGAGCAGAGAGTGCACCAATCCCAAGTGCAGAGGATGAGTCTCCA 784  
Qy 784 CTGCGCGCCCATATGTTAGAAACACCCCGCAGATTCACCAAGCCCTCTCAAGAAACCAAGC 843  
Db 785 CTGCGCGCCCATATGTTAGAAACACCCCGCAGATTCACCAAGCCCTCTCAAGAAACCAAGC 844  
Qy 844 CTGCGGCAACTCTGTTTAAAGCAGCTTGGGAGGAGGCTTACGCGCGCTCTGCTCGATCA 903  
Db 845 CTGCGGCAACTCTGTTTAAAGCAGCTTGGGAGGAGGCTTACGCGCGCTCTGCTCGATCA 904  
Qy 904 GAACTCCCAAACTGATCAGCCCTTGCATGCTGACACCGTGTGAAACTGCACAC 963  
Db 905 GAACTCCCAAACTGATCAGCCCTTGCATGCTGACACCGTGTGAAACTGCACAC 964  
Qy 964 CCCCAGGAGGAGGCCCCCTGCGCCACGACACCCCTTCCCTATAGCAGACTGCT 1023  
Db 965 CCCCAGGAGGAGGCCCCCTGCGCCACGACACCCCTTCCCTATAGCAGACTGCT 1024  
Qy 1024 CATCCCTTCCATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTGGAGTCTTCTG 1083  
Db 1025 CATCCCTTCCATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTGGAGTCTTCTG 1084  
Qy 1084 GGCATACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGACACCACTGAGCAAACTG 1143  
Db 1085 GGCATACTGGCTGTGTAGACAGCCAGAAACCTTGCCTGACACCACTGAGCAAACTG 1144  
Qy 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGCTGCTGCT 1203  
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGCTGCTGCT 1204  
Qy 1204 CGTGTGCTCATGAGAACTTTCTGTGAGAAATACCTAGTGCATGCTCTGCAAGGAGC 1263  
Db 1205 CGTGTGCTCATGAGAACTTTCTGTGAGAAATACCTAGTGCATGCTCTGCAAGGAGC 1264  
Qy 1264 GTGAGCTCAAGCCAGGCCCCAAGCTGTGACAGCTTGGCCAAAGCTGAGGAGCAGGAGC 1323  
Db 1265 GTGAGCTCAAGCCAGGCCCCAAGCTGTGACAGCTTGGCCAAAGCTGAGGAGCAGCAGG 1324  
Qy 1324 TCCAGATCCCGGAGCCAGCCCAAACTCAGGACACAGGAGTGTCTGCTCACTGAG 1383  
Db 1325 TCCAGATCCCGGAGCCAGCCCAAACTCAGGACACAGGAGTGTCTGCTCACTGAG 1384  
Qy 1384 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACCAAGCTC 1443  
Db 1385 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCACTGGGCCACGACCAAGCTC 1444  
Qy 1444 CGCTGGGACAGGCTCTCTCGGAGAGGTGACAGGATGGAGGACAAAGCAGCTGGCTTC 1503  
Db 1445 CGCTGGGACAGGCTCTCTCGGAGAGGTGACAGGATGGAGGACAAAGCAGCTGGCTTC 1504

Qy 1504 CAGTGCCTGTCAAAAAGGTCGCTTGAAGTATTTTCGGGACAGAGAGCTGATGCAATGT 1563  
Db 1505 CAGTGCCTGTCAAAAAGGTCGCTTGAAGTATTTTCGGGACAGAGAGCTGATGCAATGT 1564  
Qy 1564 GCAGGATTGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGAGGGCTTGG 1623  
Db 1565 GCAGGATTGACCTCACCCAGAAATTTGTCCTTTGTATGGAGCTGTGAGAGAGGGCTTGG 1624  
Qy 1624 GTCAACATCTTCATGGAGCTGCTGAAAGTGGCTCCCTGGGCCAGCTGCTCAAGGAGCAG 1683  
Db 1625 GTCAACATCTTCATGGAGCTGCTGAAAGTGGCTCCCTGGGCCAGCTGCTCAAGGAGCAG 1684  
Qy 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGAA 1743  
Db 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTGGGCCAGGGCCCTGGAGGGTCTGAA 1744  
Qy 1744 TACCTCCACTCAGAAAGATTTCTGATGGGAGAGCTCAAGCTGACAAAGTGTCTCTGTC 1803  
Db 1745 TACCTCCACTCAGAAAGATTTCTGATGGGAGAGCTCAAGCTGACAAAGTGTCTCTGTC 1804  
Qy 1804 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCATGCTGTGCTTCAACCTGAT 1863  
Db 1805 AGCGATGGGAGCCACGAGCCCTCTGTGACTTTTGGCCATGCTGTGCTTCAACCTGAT 1864  
Qy 1864 GGCTTGGGAAAGTCTTGTCTCAGAGGGAGCTACATCCCTGGGCACAGAGCCACATGGCT 1923  
Db 1865 GGCTTGGGAAAGTCTTGTCTCAGAGGGAGCTACATCCCTGGGCACAGAGCCACATGGCT 1924  
Qy 1924 CCGAGGTGTGTGCTGGGACAGGAGTGCAGACCCCAAGGTGGATGTCTGGAGCAGCTGTGT 1983  
Db 1925 CCGAGGTGTGTGCTGGGACAGGAGTGCAGACCCCAAGGTGGATGTCTGGAGCAGCTGTGT 1984  
Qy 1984 ATGATGTGCACATGCTCAACGGCTGCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC 2043  
Db 1985 ATGATGTGCACATGCTCAACGGCTGCACCCCTGGACTCAGTTCCTCCGAGGGCCGCTC 2044  
Qy 2044 TGCCTCAAGATTTGCGAGGAGCTCCGCTGTGAGGAGAGATCCCAACCTCTCTGGGCCCT 2103  
Db 2045 TGCCTCAAGATTTGCGAGGAGCTCCGCTGTGAGGAGAGATCCCAACCTCTCTGGGCCCT 2104  
Qy 2104 CTCAAGCCAGGCTATCCAGAGGGCTGAGGAAAGAGCCATCCACCGCTGTCTGCA 2163  
Db 2105 CTCAAGCCAGGCTATCCAGAGGGCTGAGGAAAGAGCCATCCACCGCTGTCTGCA 2164  
Qy 2164 GCGAGCTGGAGGAGGAGTGAACCGGGCTACAGCAAGTGGAGGCTCTGAAAGCCCT 2223  
Db 2165 GCGAGCTGGAGGAGGAGTGAACCGGGCTACAGCAAGTGGAGGCTCTGAAAGCCCT 2224  
Qy 2224 TGGAGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCATTAACCCAG 2283  
Db 2225 TGGAGGGAGAAATATAAGAAACCAAGACATCCACCGCAATCAAGCCATTAACCCAG 2284  
Qy 2284 ACCCTCATGCCAGCCGAGAGCTTTTCGCAAGGGCCCCCAGGGCCCCCGCCAGCTGAG 2343  
Db 2285 ACCCTCATGCCAGCCGAGAGCTTTTCGCAAGGGCCCCCAGGGCCCCCGCCAGCTGAG 2344  
Qy 2344 GAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCCAAGAGAGCCCCAGAGGCA 2403  
Db 2345 GAGCAACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCCAAGAGAGCCCCAGAGGCA 2404  
Qy 2404 AACAGTCTCTCTCTGACTTTTGGAGCAAGGAGGCTGCGGATGTGGAAACCTTACT 2463  
Db 2405 AACAGTCTCTCTCTGACTTTTGGAGCAAGGAGGCTGCGGATGTGGAAACCTTACT 2464  
Qy 2464 CTGTCTCTCTCTGGAGCCAGCCCTTGCAGAAACCCAGCTCACCCAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCTCTCTCTGGAGCCAGCCCTTGCAGAAACCCAGCTCACCCAGAGCGGAAAGCAACC 2524  
Qy 2524 GTCCCGAGCAGGAACTGACAGAGCTGGAATATTAATTAATTTCTCAACAGCCTGTCCAG 2583  
Db 2525 GTCCCGAGCAGGAACTGACAGAGCTGGAATATTAATTAATTTCTCAACAGCCTGTCCAG 2584





; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their  
; TITLE OF INVENTION: Inhibition  
; FILE REFERENCE: AXM-004.1 US  
; CURRENT APPLICATION NUMBER: US/09/981,397A  
; CURRENT FILING DATE: 2002-06-28  
; PRIOR FILING DATE: 60/240,750  
; PRIOR FILING DATE: 2000-10-16  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 17  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-981-397A-17

Query Match 99.6%; Score 4578.4; DB 45; Length 4596;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy	5	GGGGACTGTGCGGTGTGGAACGTGTAGCTGTGGA-AGGTGACCTCTTTACCAATTGAGG	63
Db	5	GGGGACTGTGCGGTGTGGAACGTGTAGCTGTGGAAGGTGACTCTCTTACCAATTGAGG	64
Qy	64	ATGTTTGGAGGATGAGTATGTGTGCGAGGACACATAAACAGCAGACACCTTTGCC	123
Db	65	ATGTTTGGAGGATGAGTATGTGTGCGAGGACACATAAACAGCAGACACCTTTGCC	124
Qy	124	CCTGCTTTCTCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT	183
Db	125	CCTGCTTTCTCCCAACCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT	184
Qy	184	GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAGCTGGGATGGCAGTGATG	243
Db	185	GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCACAAGCTGGGATGGCAGTGATG	244
Qy	244	GAATGGCTGCGCAGGTGCGCTGCGCTCAGAGTGGGGGACGAGAGAACTCCCAAG	303
Db	245	GAATGGCTGCGCAGGTGCGCTGCGCTCAGAGTGGGGGACGAGAGAACTCCCAAG	304
Qy	304	CCAAAGGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC	363
Db	305	CCAAAGGAGAGAGCGCCCACTGGGGAAGAAACAGAGCTCCGTCTACAAGCTTGAGGCC	364
Qy	364	GTGAGAGAGAGCCCTGTGTTCTGCGAAAGTGGAGATCCTGAATGACGTGATTTACCAAG	423
Db	365	GTGAGAGAGAGCCCTGTGTTCTGCGAAAGTGGAGATCCTGAATGACGTGATTTACCAAG	424
Qy	424	GGCAGAGCAAGGAGGTCCGAGGACAGGCGAGCTGCCATCTCTATCATCGCCAGGCT	483
Db	425	GGCAGAGCAAGGAGGTCCGAGGACAGGCGAGCTGCCATCTCTATCATCGCCAGGCT	484
Qy	484	GAGTGTGAGATAGCCAGAGTTGAGCCCACTTTTCAGAACGATTTTCATCGCTGGG	543
Db	485	GAGTGTGAGATAGCCAGAGTTGAGCCCACTTTTCAGAACGATTTTCATCGCTGGG	544
Qy	544	TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT	603
Db	545	TCCAAACAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATGCT	604
Qy	604	ACAGAGGCAAAATGGCCCGTGTGTTGGAAGGGAAGCGTCGAGCAAAAGCCCGGAAG	663
Db	605	ACAGAGGCAAAATGGCCCGTGTGTTGGAAGGGAAGCGTCGAGCAAAAGCCCGGAAG	664
Qy	664	AAAAGGAGAGAGAGCTCAAAATCCCTGCTCATGAGAGTGGCTTGGCCAAACCC	723
Db	665	AAAAGGAGAGAGAGCTCAAAATCCCTGCTCATGAGAGTGGCTTGGCCAAACCC	724
Qy	724	CTCCCAAGAGAGAGAGCTCAAAATCCCTGCTCATGAGAGTGGCTTGGCCAAACCC	783
Db	725	CTCCCAAGAGAGAGAGCTCAAAATCCCTGCTCATGAGAGTGGCTTGGCCAAACCC	784
Qy	784	CTCCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTCTGAAGAAACCAAGGC	843

Db	785	CTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTACCAAGCTCTCTGAAGAACCAAGGC	844
Qy	844	CTTGGGCAACTCTGTTTTTAAGCAGCTTTGGCGAGGCGCTTACGGCGGCTCTCTCGATCA	903
Db	845	CTTGGGCAACTCTGTTTTTAAGCAGCTTTGGCGAGGCGCTTACGGCGGCTCTCTCGATCA	904
Qy	904	GAACCTCCACAAACTGATCAGCCCTTGCATATGTCTGAACACACGTTGGAAACTGCACCAC	963
Db	905	GAACCTCCACAAACTGATCAGCCCTTGCATATGTCTGAACACACGTTGGAAACTGCACCAC	964
Qy	964	CCCCAGGACGAGGCGCCCTGCTGCCACGACACCCCTTCCCTATATAGCAGACTGCT	1023
Db	965	CCCCAGGACGAGGCGCCCTGCTGCCACGACACCCCTTCCCTATATAGCAGACTGCT	1024
Qy	1024	CATCCCTTCCCATTTCCACCTCTCTCAGCCCTCGAAACCTCTCCCTCTGGAGTCTTCTCTG	1083
Db	1025	CATCCCTTCCCATTTCCACCTCTCTCAGCCCTCGAAACCTCTCCCTCTGGAGTCTTCTCTG	1084
Qy	1084	GGCAAACTGGCTGTGTAGACAGCAGAAACCTTGGCTGTGACCCACACCTGAGCAACTG	1143
Db	1085	GGCAAACTGGCTGTGTAGACAGCAGAAACCTTGGCTGTGACCCACACCTGAGCAACTG	1144
Qy	1144	GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGCTGCTGTCT	1203
Db	1145	GCCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTGGAGCCAGCTGCTGTCT	1204
Qy	1204	CGTGGTCCCATAGAAAGTTTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGACG	1263
Db	1205	CGTGGTCCCATAGAAAGTTTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGACG	1264
Qy	1264	GTGAGCTCAAGCCAGGCGCCACAGCCTGACAGCTGGCGCAAGAGCTGGGACAGCGGGC	1323
Db	1265	GTGAGCTCAAGCCAGGCGCCACAGCCTGACAGCTGGCGCAAGAGCTGGGACAGCGGGC	1324
Qy	1324	TCCAGATCCCGGAGCGCCCAAAACTCAGACACAGAGGTGCTCTGCTCACTGAG	1383
Db	1325	TCCAGATCCCGGAGCGCCCAAAACTCAGACACAGAGGTGCTCTGCTCACTGAG	1384
Qy	1384	AACTCAAGCCAGTGGATTTATAGTACCGAGAAAGTCCACTGGGCGCACGCCAGCTC	1443
Db	1385	AACTCAAGCCAGTGGATTTATAGTACCGAGAAAGTCCACTGGGCGCACGCCAGCTC	1444
Qy	1444	CGCTGGGACAGAGCTCTCTCGAGAGGTGACAGGATGGAGGACAGCAGACTGGCTTC	1503
Db	1445	CGCTGGGACAGAGCTCTCTCGAGAGGTGACAGGATGGAGGACAGCAGACTGGCTTC	1504
Qy	1504	CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTGGGCGAGAGAGCTGATGCGCATGT	1563
Db	1505	CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTGGGCGAGAGAGCTGATGCGCATGT	1564
Qy	1564	GCAGATTGACCTCACCCAGAAATTTGCTTGTATGAGCTGTGAGAGAGGCGCTTGG	1623
Db	1565	GCAGATTGACCTCACCCAGAAATTTGCTTGTATGAGCTGTGAGAGAGGCGCTTGG	1624
Qy	1624	GTCAAACTCTTTCTGAGAGTGGCTTCCCTGGGCGCAGCTGGTCAAGGAGCAG	1683
Db	1625	GTCAAACTCTTTCTGAGAGTGGCTTCCCTGGGCGCAGCTGGTCAAGGAGCAG	1684
Qy	1684	GGCTGTCTCCAGAGGACCGGCGCTGTATCTA CTTGGGCGCAGGCGCTTGGAGGCTTGAA	1743
Db	1685	GGCTGTCTCCAGAGGACCGGCGCTGTATCTA CTTGGGCGCAGGCGCTTGGAGGCTTGAA	1744
Qy	1744	TACCTCCACTCAGAAAGGATTTGCTGAGGAGCGTCAAGAGTCAACAGCTCTCTGTGCC	1803
Db	1745	TACCTCCACTCAGAAAGGATTTGCTGAGGAGCGTCAAGAGTCAACAGCTCTCTGTGCC	1804
Qy	1804	AGCATGGGAGCCACGAGCGCTCTGTGATCTTTGGCCATGCTGTGTCTTCAACCTGAT	1863
Db	1805	AGCATGGGAGCCACGAGCGCTCTGTGATCTTTGGCCATGCTGTGTCTTCAACCTGAT	1864
Qy	1864	GGCTGGGAAAGTCTTTGCTCACAGGGGACTATACCTCTGGCAGAGAGCCACATGGCT	1923
Db	1865	GGCTGGGAAAGTCTTTGCTCACAGGGGACTATACCTCTGGCAGAGAGCCACATGGCT	1924

1924 CCGGAGGTGGTCTGGCGAGGAGCTGCGACGCGCAAGGTGGATGTCTTGAGCAGAGCTGCTGT 1983  
1925 CCGGAGGTGGTCTGGCGAGGAGCTGCGACGCGCAAGGTGGATGTCTTGAGCAGAGCTGCTGT 1984  
1984 ATGATGCTGCACATGCTCAACGGCTGCGCACCCCTGGACTCAGTCTTCTCCGAGGGCGGCTC 2043  
1985 ATGATGCTGCACATGCTCAACGGCTGCGCACCCCTGGACTCAGTCTTCTCCGAGGGCGGCTC 2044  
2044 TGCCTCAAGATTGCCAGGAGGCTTCGGCTGTGAGGGAGATCCACCCCTCTGCGCCGCT 2103  
2045 TGCCTCAAGATTGCCAGGAGGCTTCGGCTGTGAGGGAGATCCACCCCTCTGCGCCGCT 2104  
2104 CTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163  
2105 CTCACAGCCAGGCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164  
2164 GCGGAGCTGGGAGGAAGTGTAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223  
2165 GCGGAGCTGGGAGGAAGTGTAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2224  
2224 TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCMAATCAAGGCCAATTAACACAG 2283  
2225 TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCMAATCAAGGCCAATTAACACAG 2284  
2284 ACCCTCCATGCCACCGAGAGAGCTTTCGCCAAGGGCCCAAGGCCCGCGCCAGCTGAG 2343  
2285 ACCCTCCATGCCACCGAGAGAGCTTTCGCCAAGGGCCCAAGGCCCGCGCCAGCTGAG 2344  
2344 GAGACAAAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCCAACAGAGCCCCCGAGAGCCA 2403  
2345 GAGACAAAGGAGAGCCCTTAAGCTCCAGCCTCTCTCTCCCAACAGAGCCCCCGAGAGCCA 2404  
2404 AACAAATCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT 2463  
2405 AACAAATCTCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT 2464  
2464 CTGTCTCTCTGGAGCGAGCCCTCGCAGAAACCCAGCTCACAGAGCGGAAGCAACC 2523  
2465 CTGTCTCTCTGGAGCGAGCCCTCGCAGAAACCCAGCTCACAGAGCGGAAGCAACC 2524  
2524 GTCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGTCCCAG 2583  
2525 GTCCGGAGCAGGAACTGCAGCAGCTGGAAATAGAAATATTCTCTCAACAGCCTGTCCCAG 2584  
2584 CCAATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCCTCTCC 2643  
2585 CCAATTTCTCTGGAGGAGCAGAGCAAAATTTCTCTGTCCTCAGCATCGACAGCCTCTCC 2644  
2644 CTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGCTCGCGGACACCCCTG 2703  
2645 CTGTCCGATGACAGTGAGAGAAACCCATCAAGGGCTCTCAAGCTCGCGGACACCCCTG 2704  
2704 AGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCGAGCTGGAAACATG 2763  
2705 AGCTCAGGCGTACACTCTCTGGAGCAGCCAGGCTCGAAGCTCGAAGCTGGAAACATG 2764  
2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAATCCAA 2823  
2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGGTGTGAAATCCAA 2824  
2824 ATACAGTCTCTTAATGGTGAAACCTGCAATCCCGGAGTTCACCGGGTCAAGTGGGA 2883  
2825 ATACAGTCTCTTAATGGTGAAACCTGCAATCCCGGAGTTCACCGGGTCAAGTGGGA 2884  
2884 GACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAA 2943  
2885 GACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAA 2944  
2944 GACGGGAGCCTGTTGCTTACACATGAGAGTGCAGACTCGGGCATCGACCTGCAGTGC 3003  
2945 GACGGGAGCCTGTTGCTTACACATGAGAGTGCAGACTCGGGCATCGACCTGCAGTGC 3004

3004 ACATGGCCCTGATGGCAGCTTCGCTCGAGCTGGAGGGTCAAGCATGGCCAGCTGGAG 3063  
3005 ACATGGCCCTGATGGCAGCTTCGCTCGAGCTGGAGGGTCAAGCATGGCCAGCTGGAG 3064  
3064 AACAGGCCCTAAACCTTGCCCTCCACCGCGGCTCCACATCGCCGGAAGCAGCCTTCCTG 3123  
3065 AACAGGCCCTAAACCTTGCCCTCCACCGCGGCTCCACATCGCCGGAAGCAGCCTTCCTG 3124  
3124 CTCGGTGACCATGCTGCCCTGAAACACAGAGCTCAGCCGTTCCACAGGGGATTGCCAGCC 3183  
3125 CTCGGTGACCATGCTGCCCTGAAACACAGAGCTCAGCCGTTCCACAGGGGATTGCCAGCC 3184  
3184 CCCGGCTCAAGTGGGAAACAGAGGCCCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCC 3243  
3185 CCCGGCTCAAGTGGGAAACAGAGGCCCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATGCC 3244  
3244 TCCAGGATTTCAACCTTGAGCCCTCCACCTGCTGAAACCAATCCCGCACGTGAA 3303  
3245 TCCAGGATTTCAACCTTGAGCCCTCCACCTGCTGAAACCAATCCCGCACGTGAA 3304  
3304 GAGACAGAGGAGATGGCAGAGTTACCTGGGGAAACAAACACAGGGATCTTTTCTGCC 3363  
3305 GAGACAGAGGAGATGGCAGAGTTACCTGGGGAAACAAACACAGGGATCTTTTCTGCC 3364  
3364 CCTGCTCCAGTCCAGTTCGCTGAGCCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423  
3365 CCTGCTCCAGTCCAGTTCGCTGAGCCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424  
3424 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTTGGGCCCTCAACCTC 3483  
3425 GGAGAGCAGCTTCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTTGGGCCCTCAACCTC 3484  
3484 CAGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCAAACACAGGTTCAAGTGCAGAACCA 3543  
3485 CAGCTGCTGTGAGAGTGTCAAGTGTGTAAGGGGCCAAACACAGGTTCAAGTGCAGAACCA 3544  
3544 GGTCAAGAGTATGCCCGCCCTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTGGCCCTC 3603  
3545 GGTCAAGAGTATGCCCGCCCTAGGTTAAGGGGGCCCTCTAAACCCCTTGCTGGCCCTC 3604  
3604 ACCTGGCCAGCTACCCCTTTTGGGTGAGGGGAAAGAAATGCTTGAACCTTGGGGAAGGT 3663  
3605 ACCTGGCCAGCTACCCCTTTTGGGTGAGGGGAAAGAAATGCTTGAACCTTGGGGAAGGT 3664  
3664 CCCTGTGAGAAATACACACATTTTTCAGGTTGTGCAACAGAGTCCCTGAGTTGACCTCT 3723  
3665 CCCTGTGAGAAATACACACATTTTTCAGGTTGTGCAACAGAGTCCCTGAGTTGACCTCT 3724  
3724 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCAGACATGT 3783  
3725 GGTTCAGCAAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCCCAGACATGT 3784  
3784 GCGCCCTTTGCTGGCTTACCACTTTTCCAGAGCAGAGGCCCGGAGCCCTTTCAGGC 3843  
3785 GCGCCCTTTGCTGGCTTACCACTTTTCCAGAGCAGAGGCCCGGAGCCCTTTCAGGC 3844  
3844 CCAGCAGTCCCGGAGCTCGCTGGCAGTCCCTCATCTGTAAAGTGAAGGTGAT 3903  
3845 CCAGCAGTCCCGGAGCTCGCTGGCAGTCCCTCATCTGTAAAGTGAAGGTGAT 3904  
3904 GCAGGATATGCTGCAGAGAAACAGTCTGTGGATGGACATGATCAGTGTCAAGGAAAGCAG 3963  
3905 GCAGGATATGCTGCAGAGAAACAGTCTGTGGATGGACATGATCAGTGTCAAGGAAAGCAG 3964  
3964 CAGAGAGAGAGTCCCGGCCCGCAGCCCACTATCAGTGTCCAGCGTGTGTTCCCGCAG 4023  
3965 CAGAGAGAGAGTCCCGGCCCGCAGCCCACTATCAGTGTCCAGCGTGTGTTCCCGCAG 4024  
4024 AGCAGAGTCCAGCATCAGCTGACATCAGCTGACCTCCCTGCCCTGGCCAGAGGTTACTGC 4083  
4025 AGCAGAGTCCAGCATCAGCTGACATCAGCTGACCTCCCTGCCCTGGCCAGAGGTTACTGC 4084  
4084 CGAGGCACTTTTGCACTCTGATGACCTCAAGACATTTTCATGGTGTGCCCTCTGGCAGGCG 4143



```
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTATGGCTGCCCCCTCTGGCAGGGC 4144
Qy 4144 AGGCAGGGCAGTGAACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4203
Db 4145 AGGCAGGGCAGTGAACACTGTAGGAGCATAGCAAGCCAGGAGATGGGTGAAGGACACA 4204
Qy 4204 GTCTTGAAGTGTCCACATGTCATGTAAGTCTTCAAACTCTTCCAGATTCTCTAAGAATA 4263
Db 4205 GTCTTGAAGTGTCCACATGTCATGTAAGTCTTCAAACTCTTCCAGATTCTCTAAGAATA 4264
Qy 4264 GCACCCCTTCCCACTGCCCCAGCTTAGCCCTTAGCCCTTCTCCAGGGGAGTACTCAGGACTC 4323
Db 4265 GCACCCCTTCCCACTGCCCCAGCTTAGCCCTTCTCCAGGGGAGTACTCAGGACTC 4324
Qy 4324 ACCTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCCTCAACCTCTCTGGG 4383
Db 4325 ACCTAGCATTAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCCCTCAACCTCTCTGGG 4384
Qy 4384 CGAGGGAGCGGAGACTCCGTGGGAGAACTCATTCACATCTTGGCCAAAGACAGCCTT 4443
Db 4385 CGAGGGAGCGGAGACTCCGTGGGAGAACTCATTCACATCTTGGCCAAAGACAGCCTT 4444
Qy 4444 TGTCCAGCTGTCCACATTAAGTGTGACTGCTCCCGGGGAGAGAGCCCCGGCCCCCAGCAC 4503
Db 4445 TGTCCAGCTGTCCACATTAAGTGTGACTGCTCCCGGGGAGAGAGCCCCGGCCCCCAGCAC 4504
Qy 4504 ATAAAGAACTGCAGCTTGGTGTACTGACAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCTTGGTGTACTGACAGTCTGGGTTGTAGAGAACTCTTTGTAAAGCAA 4564
Qy 4564 TAAAGTTGGGGTGTATGATGACAAATGTTAAAAA 4595
Db 4565 TAAAGTTGGGGTGTATGATGACAAATGTTAAAAA 4596
```

## RESULT 9

```
US-10-172-118-975
; Sequence 975, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_003954
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-975
```

```
Query Match 99.6%; Score 4578.4; DB 49; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 5 GGGGAGCTGTCCCGTGTGGAACGTGTAGCTGTGA-AGGTGACTCTGTTACCAATTGAGG 63
Db 5 GGGGAGCTGTCCCGTGTGGAACGTGTAGCTGTGAGAGGTGGAAGTCTGTTTACCAATTGAGG 64
Qy 64 ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAACAGGCAGAGACCCTTTTGGC 123
```

```
Db 65 ATGTTTGAGGATGAGTATGTGTGGCAGAGGCACACATAAAACAGGCAGAGACCCTTTGCC 124
Qy 124 CCTGCTTTTCTCCCCCAACCCAGCGTGACCTGTGTTCTCCAGAGTCTGGGATTTCTAAGT 183
Db 125 CCTGCTTTTCTCCCCCAACCCAGCGTGACCTGTGTTCTCCAGAGTCTGGGATTTCTAAGT 184
Qy 184 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCAAAAGCTGGAGATGGCAGTGAATG 243
Db 185 GACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGCAAAAGCTGGAGATGGCAGTGAATG 244
Qy 244 GAAATGGCTTCCCGCAGAGTGGCTTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCAAG 303
Db 245 GAAATGGCTTCCCGCAGAGTGGCTTGGCTCAGCAGTGGGGCAGCAGAGAACTCCCCAAG 304
Qy 304 CCAAAGGAAGAACGCGCCCACTGGGGAAGAAAACAGAGCTCCGTCACAAGCTTCAGAGCC 363
Db 305 CCAAAGGAAGAACGCGCCCACTGGGGAAGAAAACAGAGCTCCGTCACAAGCTTCAGAGCC 364
Qy 364 GTGAGGAAGAGCCCTGTGTTCTCGGGAAGTGGGAGATCTTGAATGACGTGATTACCAAG 423
Db 365 GTGAGGAAGAGCCCTGTGTTCTCGGGAAGTGGGAGATCTTGAATGACGTGATTACCAAG 424
Qy 424 GGCAAGCAAGGAAGAGCTCCGAGCGAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483
Db 425 GGCAAGCAAGGAAGAGCTCCGAGCGAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484
Qy 484 GAGTGTGAGTAATACCAAGAGTTCAGGCCCACTTTTTCAGAACCATTTTTCATCGCTGGG 543
Db 485 GAGTGTGAGTAATACCAAGAGTTCAGGCCCACTTTTTCAGAACCATTTTTCATCGCTGGG 544
Qy 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCAACAAATGTGGCCCATGCT 604
Qy 604 ACAGAGGCAAAATGGCCCTGTGTTGTGGAGGAAAGCCTCGCAGCAAAAGCCCGGAAG 663
Db 605 ACAGAGGCAAAATGGCCCTGTGTTGTGGAGGAAAGCCTCGCAGCAAAAGCCCGGAAG 664
Qy 664 AACCGGAAGGAAGAGAGCTCAAAAGTCCCTGCTCATCGCAGAGTGGCTTTCGCAAAACC 723
Db 665 AACCGGAAGGAAGAGAGCTCAAAAGTCCCTGCTCATCGCAGAGTGGCTTTCGCAAAACC 724
Qy 724 CTCGCCAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCCA 783
Db 725 CTCGCCAGGACCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGGATGAGTCTCCA 784
Qy 784 CTCGCCAGGACCCCTATGTTAGAAAACACCCCGCAGTTTCAACAGCTCTGAAAGAAACAGGC 843
Db 785 CTCGCCAGGACCCCTATGTTAGAAAACACCCCGCAGTTTCAACAGCTCTGAAAGAAACAGGC 844
Qy 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGGCCCTACGGCCGGCTCTGCCCTCGATCA 903
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCAGGGCCCTACGGCCGGCTCTGCCCTCGATCA 904
Qy 904 GAACTCCACAAACTGATCAGCCCTTCAATGTCTGAACCAACAGTGTGAAACTGCACCAC 963
Db 905 GAACTCCACAAACTGATCAGCCCTTCAATGTCTGAACCAACAGTGTGAAACTGCACCAC 964
Qy 964 CCCCAGGAGGAGCCCTTCCCTTCCCAAGCAGCCCTTCCCTTATAGCAGACTGCT 1023
Db 965 CCCCAGGAGGAGCCCTTCCCTTCCCAAGCAGCCCTTCCCTTATAGCAGACTGCT 1024
Qy 1024 CATCCCTTCCCATTCACCCCTTCCAGCCCTGGGAAACCTCAACCTCTGGAGTCTTCTG 1083
Db 1025 CATCCCTTCCCATTCACCCCTTCCAGCCCTGGGAAACCTCAACCTCTGGAGTCTTCTG 1084
Qy 1084 GGCATACTGCTGTGTAGACAGCAGAAAACCTTGGCTGACCCACACACCTGAGCAAACTG 1143
Db 1085 GGCATACTGCTGTGTAGACAGCAGAAAACCTTGGCTGACCCACACACCTGAGCAAACTG 1144
Qy 1144 GCCTGTGTAGACAGTGTCAAAAGCCCTTCCCTGGCCCAACACCTTGGAGCCAGCTGCTGCT 1203
```



Qy 3364 CTTGCTCAGTCGAGTTGGCTGACCCCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3423  
Db CTTGCTCAGTCGAGTTGGCTGACCCCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424  
Qy 3424 GGAGAGCAGCTTCCAGCCTGGGTGAGAAGGGGTGGCGAGCCCTTCGSCCCTTCCACCTC 3483  
Db GGAGAGCAGCTTCCAGCCTGGGTGAGAAGGGGTGGCGAGCCCTTCGSCCCTTCCACCTC 3484  
Qy 3484 CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGGCCCAAACTCAGGTTTCAGTGCAGAACCA 3543  
Db CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGGCCCAAACTCAGGTTTCAGTGCAGAACCA 3544  
Qy 3544 GGTGAGCAGGTATGCGCGCCCGTAGGTAAAGGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3603  
Db GGTGAGCAGGTATGCGCGCCCGTAGGTAAAGGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3604  
Qy 3604 ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAAAGTGCCTGACCCCTGGGAAAGGCT 3663  
Db ACCTGGCCAGCTCACCCCTTTTGGGTGAGGGGAAAAGAAAGTGCCTGACCCCTGGGAAAGGCT 3664  
Qy 3664 CCCTGGTGAATAACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3723  
Db CCCTGGTGAATAACACCACTTTTTCAGGTTGTTGCAACACAGGTCCTGAGTTGACCTCT 3724  
Qy 3724 GGTTCAGCAAGGACCAAGAGGTGTGAAGTGAAGTGTTCAGTTCCTGAGTCCCAAGATGT 3783  
Db GGTTCAGCAAGGACCAAGAGGTGTGAAGTGAAGTGTTCAGTTCCTGAGTCCCAAGATGT 3784  
Qy 3784 GCGCCTTTGCTGGCTACCACTTTTCCCGAGAGCAGAGCCCGAGCCCTTCAGGC 3843  
Db GCGCCTTTGCTGGCTACCACTTTTCCCGAGAGCAGAGCCCGAGCCCTTCAGGC 3844  
Qy 3844 CCAGCACTGCCCGCCAGACTCGCTGGCACTTCCTCATCTGTAAGGTGAAGGGTGTAT 3903  
Db CCAGCACTGCCCGCCAGACTCGCTGGCACTTCCTCATCTGTAAGGTGAAGGGTGTAT 3904  
Qy 3904 GCAGGATATGCTGACAGGAACAGTCTGTGATGGACATGATCAGTGTCTAAGGAAAGCAG 3963  
Db GCAGGATATGCTGACAGGAACAGTCTGTGATGGACATGATCAGTGTCTAAGGAAAGCAG 3964  
Qy 3964 CAGAGAGAGCTCGGGGCCCCAGCCCCATATCAGTGTCCAGGGTCTGGTTCCTCCAG 4023  
Db CAGAGAGAGCTCGGGGCCCCAGCCCCATATCAGTGTCCAGGGTCTGGTTCCTCCAG 4024  
Qy 4024 AGCAGAGCTCAGCATCACACTGACACTCACCCCTGCCCTGCCCTGGCCAGAGGGTACTGC 4083  
Db AGCAGAGCTCAGCATCACACTGACACTCACCCCTGCCCTGGCCAGAGGGTACTGC 4084  
Qy 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGCGAGGC 4143  
Db CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTCTGCGAGGC 4144  
Qy 4144 AGGCGAGGAGTGCACACTGTAGAGCATAGCAAGCCAGGAGATGGGTGAAGGGACACA 4203  
Db AGGCGAGGAGTGCACACTGTAGAGCATAGCAAGCCAGGAGATGGGTGAAGGGACACA 4204  
Qy 4204 GTCTTTGAGCTGTCCACATGCATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGAATA 4263  
Db GTCTTTGAGCTGTCCACATGCATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGAATA 4264  
Qy 4264 GCACCCCTTCCCACTTGGCCCGAGCTTAGCCTCTTTCCAGGGGAGCTACTCAGGACTC 4323  
Db GCACCCCTTCCCACTTGGCCCGAGCTTAGCCTCTTTCCAGGGGAGCTACTCAGGACTC 4324  
Qy 4324 ACGTAGCATTTAAATCAGCTGTGAATCGTTCAGGGGGTGTCTGCTAGCTCAACCTCTCTGG 4383  
Db ACGTAGCATTTAAATCAGCTGTGAATCGTTCAGGGGGTGTCTGCTAGCTCAACCTCTCTGG 4384  
Qy 4384 GCAGGGAGCGGAGACTCCGTGGGAGAGCTCAATCCCACTCTTCCAGATCTTCCAGAGAGCCTT 4443  
Db GCAGGGAGCGGAGACTCCGTGGGAGAGCTCAATCCCACTCTTCCAGATCTTCCAGAGAGCCTT 4444  
Qy 4444 TGTCCAGCTGTCCACATTTGAGTCAGACTGTCTCCCGGGGAGAGAGCCCGGCCCCAGCAC 4503

Db 4445 TGTCCAGCTGTCCCAATTTGAGTCAGACTGTCTCCCGGGGAGAGAGCCCGGCCCCAGCAC 4504  
Qy 4504 ATAAAGAACTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAA 4563  
Db 4505 ATAAAGAACTGCAGCCTTGGTACTGACAGTCTGGGTGTAGAGAACTCTTTGTAAAGCAA 4564  
Qy 4564 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4595  
Db 4565 TAAAGTTGGGGTGATGACAAATGTTAAAAA 4596

RESULT 10  
US-10-283-975A-294  
; Sequence 294, Application US/10283975A  
; GENERAL INFORMATION:  
; APPLICANT: Ortho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia  
; FILE REFERENCE: CDS 293 PCT  
; CURRENT APPLICATION NUMBER: US/10/283,975A  
; PRIOR FILING DATE: 2002-10-30  
; PRIOR APPLICATION NUMBER: 60/340,938  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/338,997  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/340,081  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/341,012  
; PRIOR FILING DATE: 2001-10-30  
; NUMBER OF SEQ ID NOS: 900  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-283-975A-294

Query Match 99.6%; Score 4578.4; DB 51; Length 4596;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 5 GGGGAGCTGTGCGCTGTGGAAACGTGTAGCTGTGTA-AGTGGACTCTGTTACCATTGAGG 63  
Db 5 GGGGAGCTGTGCGCTGTGGAAACGTGTAGCTGTGTTAGAGGGTGGACTCTGTTACCATTGAGG 64  
Qy 64 ATGTTTGGAGATGAGTATGTGTGGCAGAGSCACATAAAGCAGAGACCTTTTCC 123  
Db 65 ATGTTTGGAGATGAGTATGTGTGGCAGAGSCACATAAAGCAGAGACCTTTTCC 124  
Qy 124 CTTGCCCTTTCTCCCAACCAAGCTGACCTGTGTTCTCCCAAGTCTGGGATTTCTAAGT 183  
Db 125 CTTGCCCTTTCTCCCAACCAAGCTGACCTGTGTTCTCCCAAGTCTGGGATTTCTAAGT 184  
Qy 184 GACCTGCTGTGTTGGTCTCTCTCAGGATGAGCAAGCTGGAGATGGCAGTGTATG 243  
Db 185 GACCTGCTGTGTTGGTCTCTCTCAGGATGAGCAAGCTGGAGATGGCAGTGTATG 244  
Qy 244 GAAATGCGCTCCCGAGTGGCTTGGCTCAGAGTGGGGCAGAGAGAACTCCCCAAG 303  
Db 245 GAAATGCGCTCCCGAGTGGCTTGGCTCAGAGTGGGGCAGAGAGAACTCCCCAAG 304  
Qy 304 CCAAGGAGAGAGCGCCGCACTGGGGAAGAAAACAGAGCTCCGTCTACAAGCTTTGAGGCC 363  
Db 305 CCAAGGAGAGAGCGCCGCACTGGGGAAGAAAACAGAGCTCCGTCTACAAGCTTTGAGGCC 364  
Qy 364 GTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCTCTGAATGACGTGATTACCAAG 423  
Db 365 GTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGAGATCTCTGAATGACGTGATTACCAAG 424  
Qy 424 GGCAAGCAGAGGAGGCTCGAGGAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 483  
Db 425 GGCAAGCAGAGGAGGCTCGAGGAGGGCCAGCTGCCATCTCTATCATCGCCAGGCT 484

QY 484 GAGTGTGAGATAGCCAGAGTTGAGCCCGACCTTTTCAGAACGATTTTCATCGCTGGG 543  
DB 485 GAGTGTGAGATAGCCAGAGTTGAGCCCGACCTTTTCAGAACGATTTTCATCGCTGGG 544  
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCATGCT 603  
DB 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCATGCT 604  
QY 604 ACAGAGGCAAAATGGCCGTGTGTGTTGGAGGGAAGCGTCGAGCAAAAGCCCGGAAG 663  
DB 605 ACAGAGGCAAAATGGCCGTGTGTGTTGGAGGGAAGCGTCGAGCAAAAGCCCGGAAG 664  
QY 664 AAACCGAAGAGAGAGTCAAGTCCCTGCTCATGAGAGTGGCCCTTGCCCAAAACC 723  
DB 665 AAACCGAAGAGAGAGTCAAGTCCCTGCTCATGAGAGTGGCCCTTGCCCAAAACC 724  
QY 724 CTCCCCAGAGCCCTTGAGCAGAGAGTGCACCATCCAGTGCAGAGGATGAGTCTCCA 783  
DB 725 CTCCCCAGAGCCCTTGAGCAGAGAGTGCACCATCCAGTGCAGAGGATGAGTCTCCA 784  
QY 784 CTCCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTGAAAGAACAGGC 843  
DB 785 CTCCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCTCTGAAAGAACAGGC 844  
QY 844 CTTCGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCTTACGCCGGCTCTGCCTCGATCA 903  
DB 845 CTTCGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCTTACGCCGGCTCTGCCTCGATCA 904  
QY 904 GAACTCCCAAAACTGATCAGCCCTTGCATGCTGAAACCAAGTGTGAAACTGCACCAAC 963  
DB 905 GAACTCCCAAAACTGATCAGCCCTTGCATGCTGAAACCAAGTGTGAAACTGCACCAAC 964  
QY 964 CCCCAGGACGAGGCCCCCTGCCCCCTGCCCCAGCCGACCTTCCCTATAGCAGACTGCT 1023  
DB 965 CCCCAGGACGAGGCCCCCTGCCCCCTGCCCCAGCCGACCTTCCCTATAGCAGACTGCT 1024  
QY 1024 CATCCCTCCCATTCACCCCTCTCAGCCCTGGAACCTCACCCCTTGAGTCTTCCTG 1083  
DB 1025 CATCCCTCCCATTCACCCCTCTCAGCCCTGGAACCTCACCCCTTGAGTCTTCCTG 1084  
QY 1084 GGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACACCTGAGCAAACTG 1143  
DB 1085 GGCNACTGGCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACACCTGAGCAAACTG 1144  
QY 1144 GCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGAGCCGAGCTGCTGTCT 1203  
DB 1145 GCTGTGTAGACAGTCCAAAGCCCTGCTGCGCCACACCTTGAGCCGAGCTGCTGTCT 1204  
QY 1204 CGTGTGCCCATGAGAAATTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1263  
DB 1205 CGTGTGCCCATGAGAAATTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGCAGC 1264  
QY 1264 GTGAGCTCAAGCCAGCCACAGCTTACAGCTTGGCCAGAACCTTGGCAGCAGCGGGC 1323  
DB 1265 GTGAGCTCAAGCCAGCCACAGCTTACAGCTTGGCCAGAACCTTGGCAGCAGCGGGC 1324  
QY 1324 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACCAAGAGGCTCTGCTCACTGAG 1383  
DB 1325 TCCAGATCCCGGAGCCAGCCCAAACTGAGGACCAAGAGGCTCTGCTCACTGAG 1384  
QY 1384 AAATCAAGCAGTGTGATATGATACCGAAGAGTCCACTTGGCCACGACCAAGCTC 1443  
DB 1385 AAATCAAGCAGTGTGATATGATACCGAAGAGTCCACTTGGCCACGACCAAGCTC 1444  
QY 1444 CGCTGGCAGAGGCTCTTCCGAGAGGTGCAAGGATGAGGACCAAGCAGACTGGCTTC 1503  
DB 1445 CGCTGGCAGAGGCTCTTCCGAGAGGTGCAAGGATGAGGACCAAGCAGACTGGCTTC 1504  
QY 1504 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTCCGGCAGAGAGCTGTGATGCT 1563  
DB 1505 CAGTGGCTGTCAAAAGGTGCGCTGGAAGTATTTCCGGCAGAGAGCTGTGATGCT 1564  
QY 1564 GCAGGATGACCTCACCCAGAAATGTCCTTTGTATGGAGCTGTGAGAGAGGGCTTGG 1623

DB 1565 GCAGGATGACCTCACCCAGAAATGTCCTTTGTATGAGAGCTGTGAGAGAGGGCTTGG 1624  
QY 1624 GTCAAACTCTTATGAGAGTGTGGAAGAGTGGCTCCCTGGCCAGCTGGTCAAGGAGCAG 1683  
DB 1625 GTCAAACTCTTATGAGAGTGTGGAAGAGTGGCTCCCTGGCCAGCTGGTCAAGGAGCAG 1684  
QY 1684 GGCTGTCTCCACAGAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGGTCTGGAA 1743  
DB 1685 GGCTGTCTCCACAGAGGACCGGGCCCTGTACTACTCTGGCCAGGCCCTGGAGGGTCTGGAA 1744  
QY 1744 TACCTCCACTCAGGAAGATTCTGCAATGGGAGCTCAAAAGCTGACAAAGTGTCTCTTCC 1803  
DB 1745 TACCTCCACTCAGGAAGATTCTGCAATGGGAGCTCAAAAGCTGACAAAGTGTCTCTTCAACCTGAT 1804  
QY 1804 AGCAGTGGGAGCAGCAGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGAT 1863  
DB 1805 AGCAGTGGGAGCAGCAGAGCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGAT 1864  
QY 1864 GGCTGGGAAGTCTTGTCTCACAGGGAGCTTACATCTCTGGCACAGAGACCAATGGCT 1923  
DB 1865 GGCTGGGAAGTCTTGTCTCACAGGGAGCTTACATCTCTGGCACAGAGACCAATGGCT 1924  
QY 1924 CCGAGGTGTGTCTGGCAGAGCTGCGACGCCAAAGTGTGATGTCTGGAGCAGCTGTGT 1983  
DB 1925 CCGAGGTGTGTCTGGCAGAGCTGCGACGCCAAAGTGTGATGTCTGGAGCAGCTGTGT 1984  
QY 1984 ATGATGTGCATGCTCAACGGCTGCAACCTTGGATCAGTCTTCCGAGGCGCTC 2043  
DB 1985 ATGATGTGCATGCTCAACGGCTGCAACCTTGGATCAGTCTTCCGAGGCGCTC 2044  
QY 2044 TGCCTCAAGATTGCGAGGAGCTTCCCTGTGAGGAGATCCCAACCTCTCTGCGCCCT 2103  
DB 2045 TGCCTCAAGATTGCGAGGAGCTTCCCTGTGAGGAGATCCCAACCTCTCTGCGCCCT 2104  
QY 2104 CTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTGCA 2163  
DB 2105 CTCACAGCCAGGCCATCCAAAGAGGGCTGAGGAAAGAGCCCATCCACCGCTGTGCA 2164  
QY 2164 CCGAGCTGGAGGAGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2223  
DB 2165 CCGAGCTGGAGGAGAGGTGAACCGGGCACTACAGCAAGTGGAGGTCTCAAGAGCCCT 2224  
QY 2224 TGGAGGAGAGATATAAGAACCAAGACATCCACCGCAATCAAGCCATTAACACAG 2283  
DB 2225 TGGAGGAGAGATATAAGAACCAAGACATCCACCGCAATCAAGCCATTAACACAG 2284  
QY 2284 ACCCTCCATGCCAGCCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCCGCCAGCTGAG 2343  
DB 2285 ACCCTCCATGCCAGCCAGAGAGCTTTGCGCAAGGGCCCCAGGGCCCCCGCCAGCTGAG 2344  
QY 2344 GAGACACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCCAGAGGCA 2403  
DB 2345 GAGACACAGGAGAGCCCTTAAGCTCCAGCTCTCTCCACAGAGCCCCAGAGGCA 2404  
QY 2404 AACAGTCTCTCTCTGACTTTGAGCAAGGAGGTCTGGGATGTGGGAAACCTTACCT 2463  
DB 2405 AACAGTCTCTCTCTGACTTTGAGCAAGGAGGTCTGGGATGTGGGAAACCTTACCT 2464  
QY 2464 CTGTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAGCAACC 2523  
DB 2465 CTGTCTCTCTGGAGCCAGCCCTGCGAGAAACCCAGCTCACAGAGCGGAAGCAACC 2524  
QY 2524 GTCCCGGAGCAGGAACTGACAGCTGGAAATAGAAATATTCTCAACAGCTGTCCAG 2583  
DB 2525 GTCCCGGAGCAGGAACTGACAGCTGGAAATAGAAATATTCTCAACAGCTGTCCAG 2584  
QY 2584 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCTGCTCTCAGCATCGACAGCTCTCC 2643  
DB 2585 CCATTTTCTCTGGAGGAGCAGGACAAATTTCTCTGCTCTCAGCATCGACAGCTCTCC 2644  
QY 2644 CTGTGCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCCGGGACACCTTC 2703



```
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 975
; LENGTH: 4596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-975

Query Match      99.6%; Score 4578.4; DB 52; Length 4596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 5 GGGGACTGTGCGGTGTGGAACTGTAGTGTGA-AGGTGACTCTGTATTACCATTTGAGG 63
Db 5 GGGGACTGTGCGGTGTGGAACTGTAGTGTGAGAGGTGACTCTGTATTACCATTTGAGG 64

QY 64 ATGTTTGGAGTATGATGTGTGTCAGAGGACACATATAACAGCAGACGACCTTTGCC 123
Db 65 ATGTTTGGAGTATGATGTGTGTCAGAGGACACATATAACAGCAGACGACCTTTGCC 124

QY 124 CTTGCTTTTCCCCCAACCCCAAGCTGACCTGTGTCTTCCAGAGTCTGGGATTTCTAAGT 183
Db 125 CTTGCTTTTCCCCCAACCCCAAGCTGACCTGTGTCTTCCAGAGTCTGGGATTTCTAAGT 184

QY 184 GACCTGCTCTGTGTTTGTCTCTCTCAGATGAGCACAAAGCTGGGAGATGSCAGTGATG 243
Db 185 GACCTGCTCTGTGTTTGTCTCTCTCAGATGAGCACAAAGCTGGGAGATGSCAGTGATG 244

QY 244 GAAATGGCTGCGCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAAGTCCCGAAG 303
Db 245 GAAATGGCTGCGCAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAAGTCCCGAAG 304

QY 304 CCAAAGGAGAAAGACCGCGCACTCGGGGAGAAACAGAGCTCCGCTCTACAAGCTTGAGGCC 363
Db 305 CCAAAGGAGAAAGACCGCGCACTCGGGGAGAAACAGAGCTCCGCTCTACAAGCTTGAGGCC 364

QY 364 GTGGAGAAAGACCTGTGTTTTCGGAAAGTGGGAGATCCTGAATGACGTGANTTACCAG 423
Db 365 GTGGAGAAAGACCTGTGTTTTCGGAAAGTGGGAGATCCTGAATGACGTGANTTACCAG 424

QY 424 GGCACAGCAAGAGAGCTCCAGGAGCGGCCAGCTGCCATCTCTATCATGCCCCAGGCT 483
Db 425 GGCACAGCAAGAGAGCTCCAGGAGCGGCCAGCTGCCATCTCTATCATGCCCCAGGCT 484

QY 484 GAGTGTGAGATAGCCAGAGTTCAAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 543
Db 485 GAGTGTGAGATAGCCAGAGTTCAAGCCCACTTTTCAGAACGCAATTTTCATCGCTGGG 544

QY 544 TCCAAACAGTACAGCAGTCCAGAGTCTTGATCATAGATCCCCAACAAATGTGGCCCATGCT 603
Db 545 TCCAAACAGTACAGCAGTCCAGAGTCTTGATCATAGATCCCCAACAAATGTGGCCCATGCT 604

QY 604 ACAGAGGGCAAAATGGCCCGTGTGTTGGAGGGAAGACGTCGACGACAAAGCCCGGAAG 663
Db 605 ACAGAGGGCAAAATGGCCCGTGTGTTGGAGGGAAGACGTCGACGACAAAGCCCGGAAG 664

QY 664 AAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCTTTGGCCAAACCC 723
Db 665 AAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCTTTGGCCAAACCC 724

QY 724 CTTCCCAGAGCCCTTGAGCAGAGAGTGTGCAACATCCCAGTGTGAGGAGATGAGTCTCA 783
Db 725 CTTCCCAGAGCCCTTGAGCAGAGAGTGTGCAACATCCCAGTGTGAGGAGATGAGTCTCA 784

QY 784 CTCGGGGCCCATATGTTAGAAACACCCCGAGTGTACCAAGCCCTTGAGGAAACAGGC 843
Db 785 CTCGGGGCCCATATGTTAGAAACACCCCGAGTGTACCAAGCCCTTGAGGAAACAGGC 844
```

```
844 CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGCCTTACGGCCGGCTCTGCTTCAATCA 903
845 CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGGCCTTACGGCCGGCTCTGCTTCAATCA 904
904 GAACTCCACAAACTGATCAGCCCTTSCAATGTCTGAACCAACGCTGTGAAACTGCACCAC 963
905 GAACTCCACAAACTGATCAGCCCTTSCAATGTCTGAACCAACGCTGTGAAACTGCACCAC 964
964 CCCAGGACGAGGCGCCCTGCTGCCACGACACCCCTTCCCTTCCCTATAGCAGACTGCT 1023
965 CCCAGGACGAGGCGCCCTGCTGCCACGACACCCCTTCCCTTCCCTATAGCAGACTGCT 1024
1024 CATCCCTTCCCAATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTCTGGAGTCTTCTG 1083
1025 CATCCCTTCCCAATTCACCCCTCTCAGCCCTTGGAAACCTCAACCTCTCTGGAGTCTTCTG 1084
1084 GGCAAACTGGGCTGTGTAGACAGCAGCAAAACCTTCTGCTGACCCACACCTGAGCAAACTG 1143
1085 GGCAAACTGGGCTGTGTAGACAGCAGCAAAACCTTCTGCTGACCCACACCTGAGCAAACTG 1144
1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGGAGCCAGCTGCTGCT 1203
1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTTGGAGCCAGCTGCTGCT 1204
1204 CGTGTGTAGACAGTCCAAAGCCCTGCTGGAGGAATACTTAGTGCATGCTTCTGCAAGGACG 1263
1205 CGTGTGTAGACAGTCCAAAGCCCTGCTGGAGGAATACTTAGTGCATGCTTCTGCAAGGACG 1264
1264 GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCCCTGGCCAAAGACCTTGGCAGCAGCGGC 1323
1265 GTGAGCTCAAGCCAGGCCCCACAGCTGACAGCCCTGGCCAAAGACCTTGGCAGCAGCGGC 1324
1324 TCCAGATCCCGGAGCCAGCCCAAAACTCAGGACAAACAGGGGTGCTCTGCTCACTGAG 1383
1325 TCCAGATCCCGGAGCCAGCCCAAAACTCAGGACAAACAGGGGTGCTCTGCTCACTGAG 1384
1384 AAATCAAGCCAGTGGATTATGAGTACCGAAGAAAGTCCACTGGGCCACCGCACAGCTC 1443
1385 AAATCAAGCCAGTGGATTATGAGTACCGAAGAAAGTCCACTGGGCCACCGCACAGCTC 1444
1444 CGCTGGGACAGAGCTCTCTCGGAGAGGTGACAGGATGGAGGACAAAGCAGCTGGCTTC 1503
1445 CGCTGGGACAGAGCTCTCTCGGAGAGGTGACAGGATGGAGGACAAAGCAGCTGGCTTC 1504
1504 CAGTGCCTGTCAAAGAGTGCCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCATGT 1563
1505 CAGTGCCTGTCAAAGAGTGCCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCATGT 1564
1564 GCAGGATTGACCTCAACCCAGAAATTTGCTTGTATGGAGCTGTGAGAAAGGGCTTGG 1623
1565 GCAGGATTGACCTCAACCCAGAAATTTGCTTGTATGGAGCTGTGAGAAAGGGCTTGG 1624
1624 GTCAATCTTCAATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG 1683
1625 GTCAATCTTCAATGGAGCTGCTGGAAGTGGCTCCCTGGGCCAGCTGGTCAAGAGCAG 1684
1684 GGTCTGTCTCCAGAGGACCGGCCCTGTACTACCTTGGGCCAGGGCCCTGGAGGCTCTGAA 1743
1685 GGTCTGTCTCCAGAGGACCGGCCCTGTACTACCTTGGGCCAGGGCCCTGGAGGCTCTGAA 1744
1744 TACTCTCACTCAAGAGGATTTGTCATGGGAGCGTCAAAGCTGACAAAGTGTCTCTGTC 1803
1745 TACTCTCACTCAAGAGGATTTGTCATGGGAGCGTCAAAGCTGACAAAGTGTCTCTGTC 1804
1804 AGCGATGGGAGCGCAGCGCCCTCTGTGACTTTGGCCATGCTGTGTGCTTCAACTGAT 1863
1805 AGCGATGGGAGCGCAGCGCCCTCTGTGACTTTGGCCATGCTGTGTGCTTCAACTGAT 1864
1864 GSCCTGGGAAAGTCTTCTCAGAGGAGTACATCTCCCTGGGCACAGAGACCCACATGGCT 1923
1865 GSCCTGGGAAAGTCTTCTCAGAGGAGTACATCTCCCTGGGCACAGAGACCCACATGGCT 1924
```



QY 1924 CCGAGGTGGTGGTGGGAGAGCTGCGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGT 1983  
Db 1925 CCGAGGTGGTGGTGGGAGAGCTGCGACGCCAAGGTGGATGTCTGGAGCAGCTGCTGT 1984  
QY 1984 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAATGTTCTTCCGAGGCGCGCTC 2043  
Db 1985 ATGATGCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAATGTTCTTCCGAGGCGCGCTC 2044  
QY 2044 TGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTCGCCCCCT 2103  
Db 2045 TGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCACCCCTCTCTCGCCCCCT 2104  
QY 2104 CTACAGCCAGGCCATTCAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163  
Db- 2105 CTACAGCCAGGCCATTCAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164  
QY 2164 GCGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2223  
Db 2165 GCGAGCTGGAGGGAAGGTGAACCGGCACTACAGCAAGTGGGAGGTCTGAAGAGCCCT 2224  
QY 2224 TGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAACAG 2283  
Db 2225 TGGAGGGGAGGAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTAACCAACAG 2284  
QY 2284 ACCCTCCATGCCAGCGAGAGAGCTTTGCGCAAGGGGCCCCAGGGCCCGGCCACAGCTGAG 2343  
Db 2285 ACCCTCCATGCCAGCGAGAGAGCTTTGCGCAAGGGGCCCCAGGGCCCGGCCACAGCTGAG 2344  
QY 2344 GAGACACGCGAGCGCCCTTAAGCTCCAGCTCCTCTCCACAGAGCCCGGCCACAGAGCCA 2403  
Db 2345 GAGACACGCGAGCGCCCTTAAGCTCCAGCTCCTCTCCACAGAGCCCGGCCACAGAGCCA 2404  
QY 2404 AACAAAGTCTCTCCCTTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTACCT 2463  
Db 2405 AACAAAGTCTCTCCCTTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTACCT 2464  
QY 2464 CTGTCCTCTCTGGAGCAGCGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2523  
Db 2465 CTGTCCTCTCTGGAGCAGCGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGCAACC 2524  
QY 2524 GTCCCGAGGAGGAATGCGAGCAGCTGGAATAGAAATTTCTCAACAGCCTGTCCAG 2583  
Db 2525 GTCCCGAGGAGGAATGCGAGCAGCTGGAATAGAAATTTCTCAACAGCCTGTCCAG 2584  
QY 2584 CCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTGCTGCTCAGCATTCGACAGCTCTCC 2643  
Db 2585 CCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTGCTGCTCAGCATTCGACAGCTCTCC 2644  
QY 2644 CTGTCGATGACAGTGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG 2703  
Db 2645 CTGTCGATGACAGTGAGAAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG 2704  
QY 2704 AGCTCAGGCTGACTCTCTGGAGCAGCAGCGCCAGGCTGGAAGCTCAGCTGGAACATG 2763  
Db 2705 AGCTCAGGCTGACTCTCTGGAGCAGCAGCGCCAGGCTGGAAGCTCAGCTGGAACATG 2764  
QY 2764 GTGCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCCAA 2823  
Db 2765 GTGCTGCGCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCCAA 2824  
QY 2824 ATACAGTCTCTTAATGTGTGAACACCTGCACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2883  
Db 2825 ATACAGTCTCTTAATGTGTGAACACCTGCACATCCGGGAGTTCCACCGGGTCAAAAGTGGGA 2884  
QY 2884 GACATCCGCACTGGCATCAGCAGCAGATCCAGCTGAGGCTTTCAGCTTGGTCAACCAA 2943  
Db 2885 GACATCCGCACTGGCATCAGCAGCAGATCCAGCTGAGGCTTTCAGCTTGGTCAACCAA 2944  
QY 2944 GACGGGAGCGTGTTCGTACGATGAGGTGCCAGCTCGGGCATCGACCTGCAGTGC 3003  
Db 2945 GACGGGAGCGTGTTCGTACGATGAGGTGCCAGCTCGGGCATCGACCTGCAGTGC 3004  
QY 3004 AACTGGCCCCCTGATGGCAGCTTCCCTGGAGCTGGAGGGGTCAAGCATGGCCAGCTGGAG 3063

Db 3005 AACTGGCCCCCTGATGGCAGCTTCCCTGGAGCTGGAGGGGTCAAGCATGGCCAGCTGGAG 3064  
QY 3064 AACAGGCGCTAAACCTCCCTCCACCGCGGCTCCACACTGCGCGGAAAGAGCGCTTCTCTG 3123  
Db 3065 AACAGGCGCTAAACCTCCCTCCACCGCGGCTCCACACTGCGCGGAAAGAGCGCTTCTCTG 3124  
QY 3124 CTGGTGCACGATGCTGCCCTGAAACACAGGGCTCAGCGCTTCCAGGGGATGCGCAGCC 3183  
Db 3125 CTGGTGCACGATGCTGCCCTGAAACACAGGGCTCAGCGCTTCCAGGGGATGCGCAGCC 3184  
QY 3184 CCGCGCTCAGCTGGGAACCGAGGCTCCAGCAGCAAGGTGGGGGCAAGCAGATGCC 3243  
Db 3185 CCGCGCTCAGCTGGGAACCGAGGCTCCAGCAGCAAGGTGGGGGCAAGCAGATGCC 3244  
QY 3244 TCCAGGATTTTCAACCTGAGCCCTGCCACCTGCTGAAACCAATCCGCCACGTGAA 3303  
Db 3245 TCCAGGATTTTCAACCTGAGCCCTGCCACCTGCTGAAACCAATCCGCCACGTGAA 3304  
QY 3304 GAGACAGAGGAGATGGCAGGAGTTACTCTGGGAAACAAACAGGATCTTTTCTGCC 3363  
Db 3305 GAGACAGAGGAGATGGCAGGAGTTACTCTGGGAAACAAACAGGATCTTTTCTGCC 3364  
QY 3364 CTTGCTCAGTGGAGTTGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG 3423  
Db 3365 CTTGCTCAGTGGAGTTGGCTGACCCGCTTGGATCAGTGACCATTTTGTGGCAGACAGG 3424  
QY 3424 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCCGCCCTCACCCCTC 3483  
Db 3425 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCCGCCCTCACCCCTC 3484  
QY 3484 CAGGCTCTCTGAGAGTGTCAAGTGTAAAGGGGCCAAACTCAGGTTTCAAGTGCAGAACCA 3543  
Db 3485 CAGGCTCTCTGAGAGTGTCAAGTGTAAAGGGGCCAAACTCAGGTTTCAAGTGCAGAACCA 3544  
QY 3544 GGTTCAGCAGTATGCCCGCCGCTAGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3603  
Db 3545 GGTTCAGCAGTATGCCCGCCGCTAGGTTAAGGGGCCCTCTAAACCCCTTGCCTGGCCCTC 3604  
QY 3604 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGATGCTGACCCCTGGGAAAGGCT 3663  
Db 3605 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGATGCTGACCCCTGGGAAAGGCT 3664  
QY 3664 CCCTGTAGTAATACACACACTTTTTCAGGTTGTGTAACACACAGGCTCTGAGTTGACCTCT 3723  
Db 3665 CCCTGTAGTAATACACACACTTTTTCAGGTTGTGTAACACACAGGCTCTGAGTTGACCTCT 3724  
QY 3724 GGTTCAGCCAAAGGACCAAGAGGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3783  
Db 3725 GGTTCAGCCAAAGGACCAAGAGGTGTAAAGTGAAGTGGTTCTCAGTCCCCAGACATGT 3784  
QY 3784 GCCCTTTGCTGTGGCTACCTCTTCCAGAGCAGAGGCCCCCGAGCCCTTTCAGGC 3843  
Db 3785 GCCCTTTGCTGTGGCTACCTCTTCCAGAGCAGAGGCCCCCGAGCCCTTTCAGGC 3844  
QY 3844 CCAGCACTGCCCGCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGTAT 3903  
Db 3845 CCAGCACTGCCCGCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGGTGTAT 3904  
QY 3904 GCAGGATATGCCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGCTTAAGGAAAGCAG 3963  
Db 3905 GCAGGATATGCCCTGACAGGAACAGTCTGTGGATGGACATGATCAGTGCTTAAGGAAAGCAG 3964  
QY 3964 CAGAGAGAGCTCCGGCGGCCAGGCCCATCATCAGTGTCCAGCGTGGTGTCCCCAG 4023  
Db 3965 CAGAGAGAGCTCCGGCGGCCAGGCCCATCATCAGTGTCCAGCGTGGTGTCCCCAG 4024  
QY 4024 AGCAGACTCAGCATCACACTGACACTCAGCTGCCCTGCCCTTCCGCCAGAGGCTACTGC 4083  
Db 4025 AGCAGACTCAGCATCACACTGACACTCAGCTGCCCTGCCCTTCCGCCAGAGGCTACTGC 4084  
QY 4084 CGACGGCACTTTCAGTCTGATGACCTCAAGCACTTTTCATGCTGCTTCCCTTGCAGGGC 4143

Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGACACTTTTCATGGCTGCCCCCTCTGGCAGGCG 4144  
QY 4144 AGGGCAGGCACTGACACATGTAGGACATAGCAAGCCAGAGATGGGGTGAAGGACACA 4203  
Db 4145 AGGGCAGGCACTGACACATGTAGGACATAGCAAGCCAGAGATGGGGTGAAGGACACA 4204  
QY 4204 GTCTTGAGCTGTCCACATGCATGTGACTCCTCAAACCTCTTCCAGATTCTCTAAGAATA 4263  
Db 4205 GTCTTGAGCTGTCCACATGCATGTGACTCCTCAAACCTCTTCCAGATTCTCTAAGAATA 4264  
QY 4264 GCACCCCTTCCCATTCGCCACAGCTTAGCCTCTTCTCCAGGGAGACTACTCAGACTC 4323  
Db 4265 GCACCCCTTCCCATTCGCCACAGCTTAGCCTCTTCTCCAGGGAGACTACTCAGACTC 4324  
QY 4324 ACGTAGCATTAATACAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCCTGGG 4383  
Db 4325 AGTAGCATTAATACAGCTGTGAATCGTCAGGGGTGTCTGTAGCTCAACCTCCTGGG 4384  
QY 4384 GCAGGGACCGCAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAAAGACGCTT 4443  
Db 4385 GCAGGGACCGCAGACTCCGTGGGAGAGCTCATTTCCACATCTTGCCAAAGACGCTT 4444  
QY 4444 TGTCCAGCTGTCCACATTTAGCTCAGACTGCTCCGGGGAGAGAGCCCGGCCCCAGCAGC 4503  
Db 4445 TGTCCAGCTGTCCACATTTAGCTCAGACTGCTCCGGGGAGAGAGCCCGGCCCCAGCAGC 4504  
QY 4504 ATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGTGTGTAGAGAACTCTTTGTAAAGCAA 4563  
Db 4505 ATAAAGAACTGCAGCCTTGGTACTGCAGAGTCTGGTGTGTAGAGAACTCTTTGTAAAGCAA 4564  
QY 4564 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4595  
Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

## RESULT 12

US-60-474-850-144  
; Sequence 144, Application US/60474850  
; GENERAL INFORMATION:  
; APPLICANT: Wohlgemuth, Jay  
; APPLICANT: Ly, Ngoc  
; APPLICANT: Prentice, James  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING  
; TITLE OF INVENTION: RENAL TRANSPLANT REJECTION  
; FILE REFERENCE: 506613000700  
; CURRENT APPLICATION NUMBER: US/60/474,850  
; CURRENT FILING DATE: 2003-05-30  
; NUMBER OF SEQ ID NOS: 552  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 144  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-474-850-144

Query Match 99.6%; Score 4578.4; DB 117; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 5 GGGGACTGTGCGGTGTGAACGTGTAGCTGTGTA-AGGTGACTCTGTGTACCATTTGAGG 63  
Db 5 GGGGACTGTGCGGTGTGAACGTGTAGCTGTGTAAGGTGAGG 64  
QY 64 ATGTTTGGAGGATGATGTGTGTCAGAGGCACATAAACAGCAGAGACCTTTGCC 123  
Db 65 ATGTTTGGAGGATGATGTGTGTCAGAGGCACATAAACAGCAGAGACCTTTGCC 124  
QY 124 CTTGCTTTTCTCCCAACCCCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 183  
Db 125 CTTGCTTTTCTCCCAACCCCAAGGCTGACCTGTGTTCTCCAGGCTCTGGGATTTCTAAGT 184  
QY 184 GACCTGCTCTGTGTTTGTGTTCTCTCAGGATGACACAGCCTGGGAGATGGCAGTGATG 243

Db 185 GACCTGCTCTGTGTTTGTGTTCTCTCAGGATGAGCACAAAGCTGGAGATGGCAGTGATG 244  
QY 244 GAAATGGGCTTCCCAAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCAG 303  
Db 245 GAAATGGGCTTCCCAAGGTGCCCCCTGGCTCAGCAGTGGGGCAGCAGAAAGAACTCCCCAG 304  
QY 304 CCAAAAGGAGAAAGACCCGCCCACTGGGGAGAAACAGAGCTCCGTCTCAAAAGCTTGAGGCC 363  
Db 305 CCAAAAGGAGAAAGACCCGCCCACTGGGGAGAAACAGAGCTCCGTCTCAAAAGCTTGAGGCC 364  
QY 364 GTGGAGAAAGACCTGTGTTCTCGGAAAGTGGGAGATCTTGAATGACGTGATTAACAG 423  
Db 365 GTGGAGAAAGACCTGTGTTCTGCGGAAAGTGGGAGATCTTGAATGACGTGATTAACAG 424  
QY 424 GGCACAGCAAGGAAGGCTCCGAGCAGGGCCAGCTGCTATCTATCATCGCCAGGCT 483  
Db 425 GGCACAGCAAGGAAGGCTCCGAGCAGGGCCAGCTGCTATCTATCATCGCCAGGCT 484  
QY 484 GAGTGTGAGAAATAGCAAGAGTTTCAGCCCCACCTTTTTCAGAAACGCAATTTTCATCGCTGG 543  
Db 485 GAGTGTGAGAAATAGCAAGAGTTTCAGCCCCACCTTTTTCAGAAACGCAATTTTCATCGCTGG 544  
QY 544 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCCATGCT 603  
Db 545 TCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAATGTGGCCCCATGCT 604  
QY 604 ACAGAGGGCAAAATGCGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCCGGAAG 663  
Db 605 ACAGAGGGCAAAATGCGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCCGGAAG 664  
QY 664 AAACGGAAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGAGGAGTGGCCCTTGSCCAAACCC 723  
Db 665 AAACGGAAAGAAAGAGAGCTCAAAGTCCCTGGCTCATGAGGAGTGGCCCTTGSCCAAACCC 724  
QY 724 CTCCCCAGAGACCCCTGAGCAGAGAGTGCACATCCCAAGTGCAGGAGAGATGAGTCTCA 783  
Db 725 CTCCCCAGAGACCCCTGAGCAGAGAGTGCACATCCCAAGTGCAGGAGAGATGAGTCTCA 784  
QY 784 CTGGCGCCCCCATATGTTAGAAACACCCCGCAGTTTCAACAAAGCTCTCAAGAAACAGGC 843  
Db 785 CTGGCGCCCCCATATGTTAGAAACACCCCGCAGTTTCAACAAAGCTCTCAAGAAACAGGC 844  
QY 844 CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCGCTTACGGCGGCTCTGCTCGATCA 903  
Db 845 CTTGGGCAACTCTGTGTTTAAAGCAGCTTGGCGAGGCGCTTACGGCGGCTCTGCTCGATCA 904  
QY 904 GAACTCCAAACATGATCAGCCCTTGCATGTCTGAACCAAGCTGTGAAACTGCACAC 963  
Db 905 GAACTCCAAACATGATCAGCCCTTGCATGTCTGAACCAAGCTGTGAAACTGCACAC 964  
QY 964 CCCCAGGACGGAGGCCCTGCGCCCTGCGCCACGACCCCTTCCCCCTATAGCAGACTGCT 1023  
Db 965 CCCCAGGACGGAGGCCCTGCGCCCTGCGCCACGACCCCTTCCCCCTATAGCAGACTGCT 1024  
QY 1024 CATCCCTTCCATTTCCACCTCTCTCAGCCCTTGGAAACCTCAACCTCTCGAGTCTTCTG 1083  
Db 1025 CATCCCTTCCATTTCCACCTCTCTCAGCCCTTGGAAACCTCAACCTCTCGAGTCTTCTG 1084  
QY 1084 GGCAAACTGGGCTGTGTAGACAGCCAGAAACCTTGGCTGTGACCCCACTCATGAGCAACTG 1143  
Db 1085 GGCAAACTGGGCTGTGTAGACAGCCAGAAACCTTGGCTGTGACCCCACTCATGAGCAACTG 1144  
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCCGAGCTGCTGCT 1203  
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCCAACCTTGGAGCCCGAGCTGCTGCT 1204  
QY 1204 CCGTGTGCCATGAGAAAGTCTTGTGTGAGAAATACCTAGTGCAATGCTCTGCAAGGACG 1263  
Db 1205 CCGTGTGCCATGAGAAAGTCTTGTGTGAGAAATACCTAGTGCAATGCTCTGCAAGGACG 1264  
QY 1264 GTGAGCTCAAGCCAGGCCCAAGCTGACAGCTGGCCAGAGACTGGGAGCAGCGGGC 1323  
Db 1265 GTGAGCTCAAGCCAGGCCCAAGCTGACAGCTGGCCAGAGACTGGGAGCAGCGGGC 1324



Qy	1324	TCCAGATCCCGGGAGCCCGACCCCAAAA	CTGAGACAAACGAGGGTGTCCTGCTCACTGAG	1383
Db	1325	TCCAGATCCCGGGAGCCCGACCCCAAACT	GAAGCAACGAGGGTGTCCTGCTCACTGAG	1384
Qy	1384	AAACTCAAGCCAGTGGATTATGAGTACCGAGAGAAGT	TCACCTGGGCCACGACCCAGCTC	1443
Db	1385	AAACTCAAGCCAGTGGATTATGAGTACCGAGAGAAGT	TCACCTGGGCCACGACCCAGCTC	1444
Qy	1444	CGCTGGCCAGAGGCTCCTTCGGAGAGGTGCACAGGAT	TGAGGACAAGCAGACACTGGCTTC	1503
Db	1445	CGCTGGCCAGAGGCTCCTTCGGAGAGGTGCACAGGAT	TGAGGACAAGCAGACACTGGCTTC	1504
Qy	1504	CAGTGCCTGTCAAAAAGGTGCCTTCGGAAGTATTT	CGGGCAGAGGAGCTGATGGCATGT	1563
Db	1505	CAGTGCCTGTCAAAAAGGTGCCTTCGGAAGTATTT	CGGGCAGAGGAGCTGATGGCATGT	1564
Qy	1564	GCAGGATTGACCTCACCCAGAAATGTCTCCCTTT	TGTATGGAGCTGTGAGAGAAGGCTTGG	1623
Db	1565	GCAGGATTGACCTCACCCAGAAATGTCTCCCTTT	TGTATGGAGCTGTGAGAGAAGGCTTGG	1624
Qy	1624	GTCAACATCTTCATGGAGCTGCTGGAAAGTGGCT	CCCTGGGCCAGCTGGTCAAGGAGCAG	1683
Db	1625	GTCAACATCTTCATGGAGCTGCTGGAAAGTGGCT	CCCTGGGCCAGCTGGTCAAGGAGCAG	1684
Qy	1684	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCT	TGGGCCAGGCCCCTGGAGGGTCTGGAA	1743
Db	1685	GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCT	TGGGCCAGGCCCCTGGAGGGTCTGGAA	1744
Qy	1744	TACCTCACTCAAGAAAGTTCATGATGGGGAAGT	CTAAAGCTGACAACTGCTCTGTGTC	1803
Db	1745	TACCTCACTCAAGAAAGTTCATGATGGGGAAGT	CTAAAGCTGACAACTGCTCTGTGTC	1804
Qy	1804	AGCGATGGGAGCCACGAGCCCTCTGTGACTTT	GGGCCATGCTGTGTCTTCAACCTGAT	1863
Db	1805	AGCGATGGGAGCCACGAGCCCTCTGTGACTTT	GGGCCATGCTGTGTCTTCAACCTGAT	1864
Qy	1864	GGCTGGGAAAGTCTCTTGCTCACAGGGGACT	TACATCCCTGGCCACAGAGACCCATGGCT	1923
Db	1865	GGCTGGGAAAGTCTCTTGCTCACAGGGGACT	TACATCCCTGGCCACAGAGACCCATGGCT	1924
Qy	1924	CCGAGGTGGTGTGGCCAGGAGCTGCGACGCC	CAAGGTGGATGTCTGGAGCAGCTGCTGT	1983
Db	1925	CCGAGGTGGTGTGGCCAGGAGCTGCGACGCC	CAAGGTGGATGTCTGGAGCAGCTGCTGT	1984
Qy	1984	ATGATGCTGCACATGCTCAAGGCTGCCACCC	CTGTGACTCAGTTCCTTCAGAGGCGCGCTC	2043
Db	1985	ATGATGCTGCACATGCTCAAGGCTGCCACCC	CTGTGACTCAGTTCCTTCOGAGGCGCGCTC	2044
Qy	2044	TGCCTCAAGATTGCGACGAGCCCTCCGCTGT	TGAGGGAGATCCACCCCTCTGGGCCCTT	2103
Db	2045	TGCCTCAAGATTGCGACGAGCCCTCCGCTGT	TGAGGGAGATCCACCCCTCTGGGCCCTT	2104
Qy	2104	CTCACGCCCGGACCATCAAGAGGGGCTGAG	GAAAGAGGCCATCCACCGCGTGTCTGCA	2163
Db	2105	CTCACGCCCGGACCATCAAGAGGGGCTGAG	GAAAGAGGCCATCCACCGCGTGTCTGCA	2164
Qy	2164	GCGAGCTGGGAGGGAAAGGTGAAACCGGG	CACTACAGCAAGTGGGAGGTCTGAAGAGCCCT	2223
Db	2165	GCGAGCTGGGAGGGAAAGGTGAAACCGGG	CACTACAGCAAGTGGGAGGTCTGAAGAGCCCT	2224
Qy	2224	TGGAGGGGAGNATATAAGNACCAAGACATCC	ACCGCCCAATCAAGCCCAATTAACCAACAG	2283
Db	2225	TGGAGGGGAGNATATAAGNACCAAGACATCC	ACCGCCCAATCAAGCCCAATTAACCAACAG	2284
Qy	2284	ACCTCTCATGCCAGCCGAGAGAGCTTTT	CGCCAAAGGGCCCCCAGGGCCCCCGGCGACTGAG	2343
Db	2285	ACCTCTCATGCCAGCCGAGAGAGCTTTT	CGCCAAAGGGCCCCCAGGGCCCCCGGCGACTGAG	2344
Qy	2344	GAGACAACAGGCAGAGCCCTTAAGCTCCAG	CTCCTCTCTCCCAACAGAGCCCCCAGAGGCCA	2403
Db	2345	GAGACAACAGGCAGAGCCCTTAAGCTCCAG	CTCCTCTCTCCCAACAGAGCCCCCAGAGGCCA	2404

Qy	2404	AA	CAAGTCTCCTCCCTTGACTTTGAGGACAGGAGGTCTGGGATGTGGGAACCCCTTACCT	2461
Db	2405	AA	CAAGTCTCCTCCCTTGACTTTGAGCAAGGAGGAGTCTGGGATGTGGGAACCCCTTACCT	2464
Qy	2464	CT	GTCTCCTCCCTGAGCAGCCCTGCGAGAAACCCAGAGTCCACAGAGCGGAAGCAACC	2523
Db	2465	CT	GTCTCCTCCCTGAGCGAGCCCTGCGAGAAACCCAGAGTCCACAGAGCGGAAGCAACC	2524
Qy	2524	GT	CCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAATATTCTCTCAACAGCCTGTCCCAG	2583
Db	2525	GT	CCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAATATTCTCTCAACAGCCTGTCCCAG	2584
Qy	2584	CA	ATTTTCTCTGAGGAGCAGGACAAATTTCTCTCGTCCCTCAGCATCGACAGCCTCTCC	2643
Db	2585	CA	ATTTTCTCTGAGGAGCAGGACAAATTTCTCTCGTCCCTCAGCATCGACAGCCTCTCC	2644
Qy	2644	CT	GTCCGATCAGCTGAGAAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG	2703
Db	2645	CT	GTCCGATCAGCTGAGAAAGAACCCATCAAGGCCCTCTCAAGCTCGCGGACACCCCTG	2704
Qy	2704	AG	CTCAGGCGTACACTCTTGAGCAGCCAGGCCGAGGCTGGAAGCTCCAGCTCGAAACATG	2763
Db	2705	AG	CTCAGGCGTACACTCTTGAGCAGCCAGGCCGAGGCTGGAAGCTCCAGCTCGAAACATG	2764
Qy	2764	GT	GTGCCCGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGAAGTCCAA	2823
Db	2765	GT	GTGCCCGGGGGGGGCCACCGACACCCCAAGCTATTTCAATGGTGAAGTCCAA	2824
Qy	2824	AT	CAGTCTCTTAATGGTGAACACTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGA	2883
Db	2825	AT	CAGTCTCTTAATGGTGAACACTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGA	2884
Qy	2884	GAC	ATGCCCACTGGCATCAGCAGCCAGATCCCACTGCAGCCTTCAGCTTGGTCAACAAA	2943
Db	2885	GAC	ATGCCCACTGGCATCAGCAGCCAGATCCCACTGCAGCCTTCAGCTTGGTCAACAAA	2944
Qy	2944	GAC	GGGAGCCTGTTCGCTACGACATGAGAGGTGCCAGATCTCGGGCATCGACCTGCAGTGC	3003
Db	2945	GAC	GGGAGCCTGTTCGCTACGACATGAGAGGTGCCAGATCTCGGGCATCGACCTGCAGTGC	3004
Qy	3004	ACA	TGGCCCTCATGTCAGCTTCGCTCGAGCTGAGGGTCAAGCATGSCCAGCTGGAG	3063
Db	3005	ACA	TGGCCCTCATGTCAGCTTCGCTCGAGCTGAGGGTCAAGCATGSCCAGCTGGAG	3064
Qy	3064	AA	CAGGCCCTAAACCTTGCCCTCCACCGCCGGCTCCACACTGCCGGAAGCAGAGCTTCCCTG	3123
Db	3065	AA	CAGGCCCTAAACCTTGCCCTCCACCGCCGGCTCCACACTGCCGGAAGCAGAGCTTCCCTG	3124
Qy	3124	CT	GGTGCAAGATGTGCTCCCTGAAAAACACAGGCTCAGCCGTTCCAGGGGATTTCCAGCC	3183
Db	3125	CT	GGTGCAAGATGTGCTCCCTGAAAAACACAGGCTCAGCCGTTCCAGGGGATTTCCAGCC	3184
Qy	3184	CCCC	GGTCAAGTGGGAACAGGGCTCGCAGCAGCAGGTGGGGCAGCAGAGATGCC	3243
Db	3185	CCCC	GGTCAAGTGGGAACAGGGCTCGCAGCAGCAGGTGGGGCAGCAGAGATGCC	3244
Qy	3244	TCC	CAGATTTACACTGAGCCCTGCCCCACACCTGTCTGAAAAAATCATCGCCACGTTGAA	3303
Db	3245	TCC	CAGATTTACACTGAGCCCTGCCCCACACCTGTCTGAAAAAATCATCGCCACGTTGAA	3304
Qy	3304	GAG	ACAAGAGGATGGCAGGATTAACCTGGGGAAACAAAAACAGGATCTTTTTCTGCC	3363
Db	3305	GAG	ACAAGAGGATGGCAGGATTAACCTGGGGAAACAAAAACAGGATCTTTTTCTGCC	3364
Qy	3364	CT	GCTCAGTTCAGTTGGCCTGACCGGCTTGATCAGTGACCATTTGTTGGCAGACAGG	3423
Db	3365	CT	GCTCAGTTCAGTTGGCCTGACCGGCTTGATCAGTGACCATTTGTTGGCAGACAGG	3424
Qy	3424	GG	GAGCAGCTCCAGCTCGGTGAGAAGGGGTGGCGAGCCCTCTCGGCCCTCACCCCTC	3483
Db	3425	GG	GAGCAGCTTCAGGCTCGGGTCAGAGGGGTGGCGAGCCCTCTCGGCCCTCACCCCTC	3484
Qy	3484	CAG	GCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGGCCCAAACTCAGGTTCACTGAGAACCA	3543

Db	3485		CAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCCAAACTTCAGGTTTCAGTGCAGAACCA	3544
Qy	3544		GGTCAGCAGGATATGCCCGCCCGTAGGTTAAAGGGGGCCCTCTAAACCCCTTGCTGGCGCTC	3603
Db	3545		GGTCAGCAGGATATGCCCGCCCGTAGGTTAAAGGGGGCCCTCTAAACCCCTTGCTGGCGCTC	3604
Qy	3604		ACCTGGCGAGCTCAACCCCTTTTGGGTGTAGGGGAAAAAGATGCTGACCTGGGGAAGGCT	3663
Db	3605		ACCTGGCGAGCTCAACCCCTTTTGGGTGTAGGGGAAAAAGATGCTGACCTGGGGAAGGCT	3664
Qy	3664		CCCTGGTAGAATACACACACTTTTCAGGTGTTTGCAACACAGGTCCTGAGTTGACCTCT	3723
Db	3665		CCCTGGTAGAATACACACACTTTTCAGGTGTTTGCAACACAGGTCCTGAGTTGACCTCT	3724
Qy	3724		GGTTTCAGCAAGGACCAAGAAAGTGTGAAGTCAAGTGGTTCCTCAGTCCCCAGACATGT	3783
Db	3725		GGTTTCAGCAAGGACCAAGAAAGTGTGAAGTCAAGTGGTTCCTCAGTCCCCAGACATGT	3784
Qy	3784		GCCCTTTGCTGTGGCTACCACTCTTTCCCGAGACAGCGGCCCGCCAGGCCCTTCAGGC	3843
Db	3785		GCCCTTTGCTGTGGCTACCACTCTTTCCCGAGACAGCGGCCCGCCAGGCCCTTCAGGC	3844
Qy	3844		CCAGCACTGCCCGAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGGTGAT	3903
Db	3845		CCAGCACTGCCCGAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAGGTGAAGGGTGAT	3904
Qy	3904		GCAGGATATGCTTCACAGGAACAGCTGTGTGATGGACATCATCAGTCTTAAGGAAGCAG	3963
Db	3905		GCAGGATATGCTTCACAGGAACAGCTGTGTGATGGACATCATCAGTCTTAAGGAAGCAG	3964
Qy	3964		CAGAGAGAGCTGCCGGGCCCCAGCCCCACTATCAGTGTCCAGCGTGTCTGGTTCCCCAG	4023
Db	3965		CAGAGAGAGCTGCCGGGCCCCAGCCCCACTATCAGTGTCCAGCGTGTCTGGTTCCCCAG	4024
Qy	4024		AGCAAGCTCAGCATCACTGAACATCAACCTGCTGCCCTTGCCCTGGCCAGAGGGTACTGC	4083
Db	4025		AGCAAGCTCAGCATCACTGAACATCAACCTGCTGCCCTTGCCCTGGCCAGAGGGTACTGC	4084
Qy	4084		CGACGGCACTTTGCACCTCTGATGACCTCAAGACACTTTTCATGGCTGCCCTCTCGCAGGGC	4143
Db	4085		CGACGGCACTTTGCACCTCTGATGACCTCAAGACACTTTTCATGGCTGCCCTCTCGCAGGGC	4144
Qy	4144		AGGCGAGGGCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4203
Db	4145		AGGCGAGGGCAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGGACACA	4204
Qy	4204		GTCTTGAGCTGTCCACATGCAATGATCTCTTCAAACTCTTCCAGATTTCTCTAAGAATA	4263
Db	4205		GTCTTGAGCTGTCCACATGCAATGATCTCTTCAAACTCTTCCAGATTTCTCTAAGAATA	4264
Qy	4264		GCACCCCTTCCCATTTGCCCGCCAGTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC	4323
Db	4265		GCACCCCTTCCCATTTGCCCGCCAGTTAGCCTCTTCTCCAGGGGAGCTACTCAGGACTC	4324
Qy	4324		ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCCTCAACCTCTCTGGG	4383
Db	4325		ACGTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCCTCAACCTCTCTGGG	4384
Qy	4384		GCAGGGACGCGGAGACTCCGTGGGGAAGCTCATTTCCCATCTTTTGCCCAAGACAGCCTT	4443
Db	4385		GCAGGGACGCGGAGACTCCGTGGGGAAGCTCATTTCCCATCTTTTGCCCAAGACAGCCTT	4444
Qy	4444		TGTCAGCTGTCCACATTTGAGTCAGACTGCTCCCGGGGAGAGACCCCGGCCGCCAGCAC	4503
Db	4445		TGTCAGCTGTCCACATTTGAGTCAGACTGCTCCCGGGGAGAGACCCCGGCCGCCAGCAC	4504
Qy	4504		ATAAAGAACTGCAGCCTTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAGCAA	4563
Db	4505		ATAAAGAACTGCAGCCTTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTTTGTAGCAA	4564
Qy	4564		TAAAGTTTGGGGTGATCAAAATGTTAAAAA	4595

Db 4565 TAAAGTTTGGGTGATGACAAATGTTAAAAA 4596

RESULT 13

US-09-949-003C-1291

; Sequence 1291, Application US/09949003C

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH NEUROPSYCH

; TITLE OF INVENTION: DISORDERS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL000791

; CURRENT APPLICATION NUMBER: US/09/949,003C

; CURRENT FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: 60/231,446

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 74065

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1291

; LENGTH: 4594

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-949-003C-1291

Query Match 96.6%; Score 4440; DB 41; Length 4584;

Best Local Similarity 99.1%; Pred. No. 0;

Matches 4549; Conservative 0; Mismatches 30; Indels 11; Gaps 8

Qy	3	CGGGGGGACCTGTCGGCTGTGGAACTGTAGCTGTGTGA-AGGTGGACTCTGTACCATTGA	61
Db	3	CGGGGGGACCTGTCGGCTGTGGAACTGTAGCTGTGTGAAGAGTGGACTCTGTGTACCATTGA	62
Qy	62	GGATGTTTGGAGTAGATATGTGTGGCAGAGGCACACATAAACAGGCAGACACCTTTTG	121
Db	63	GGATGTTTGGAGTAGATATGTGTGGCAGAGGCACACATAAACAGGCAGACACCTTTTG	122
Qy	122	CCCTGCTCTTCTCCCCCAACCCCAAGGCTGACTGTGTTCCTCCAGGCTCTGGGATTCCTAA	181
Db	123	CCCTGCTCTTCTCCCCCAACCCCAAGGCTGACTGTGTTCCTCCAGGCTCTGGGATTCCTAA	182
Qy	182	GTGACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACACAGCTGCGAGATGCGCAGTGA	241
Db	183	GTGACCTGCTCTGTGTTTGGTCTCTCTCAGATGAGCACACAGCTGCGAGATGCGCAGTGA	242
Qy	242	TGGAATATGGCTGCCCAGGTGCCCTTGCTCAGCAGTGGGGCAGCAGAGGAATCTCCCCA	301
Db	243	TGGAATATGGCTGCCCAGGTGCCCTTGCTCAGCAGTGGGGCAGCAGAGGAATCTCCCCA	302
Qy	302	AGCCAAAGGAGAAAGACGCCGCCACTGGGGAAAGAAACAGAGTCCGTCTCAACGTTGAGG	361
Db	303	AAGCCAAAGGAGAAAGACGCCGCCACTGGGGAAAGAAACAGAGTCCGTCTCAACGTTGAGG	362
Qy	362	CCGTGGAGAGAGCCCTGTGTTCTTCGGAAGTGGGAGATCCTGAATCAGCTGATACCA	421
Db	363	CCGTGGAGAGAGCCCTGTGTTCTTCGGAAGTGGGAGATCCTGAATCAGCTGATACCA	422
Qy	422	AGGSCACAGCAAGGAAGGCTCCGAGGCAGGCGCAGCTGCCATCTCTATCATCGCCCGAG	481
Db	423	AGGSCACAGCAAGGAAGGCTCCGAGGCAGGCGCAGCTGCCATCTCTATCATCGCCCGAG	482
Qy	482	CTGAGTGTGAGAAATAGCCAAAGAGTTTACGCCCCACCTTTTTCAGAACGCATTTTTCATCGCTG	541
Db	483	CTGAGTGTGAGAAATAGCCAAAGAGTTTACGCCCCACCTTTTTCAGAACGCATTTTTCATCGCTG	542
Qy	542	GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGGCCATG	601
Db	543	GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGGGCCATG	602
Qy	602	CTACAGAGGGCAAAATGGCCCGCTGTGTGTGGAAGGGGAAAGCGTCGACGAAAGCCCGGA	661
Db	603	CTACAGAGGGCAAAATGGCCCGCTGTGTGTGGAAGGGGAAAGCGTCGACGAAAGCCCGGA	662
Qy	662	AGAAACGGGAAGGAAGAGCTCAAAGTCCCTGGCTCATGACAGAGTGGCCTTTGGCCAAAC	721

Db 663 AGAAACGAGAGAGAGCTCAAAAGTCCCTGGCTCATGACGAGAGTGGCCCTTGCCCAAC 722  
Qy 722 CCCTCCCGAGACCCCTGACGAGGAGCTGCACATCCCAAGTGCAGGAGATGAGTCTC 781  
Db 723 CCCTCCCGAGACCCCTGACGAGGAGAGCTGCACATCCCAAGTGCAGGAGATGAGTCTC 782  
Qy 782 CACTCGGGCCCCATATGTTTAAAGACACCCCGCAGTTTCAACAGCCTCTGAAGAACAG 841  
Db 783 CACTCGGGCCCCATATGTTTAAAGACACCCCGCAGTTTCAACAGCCTCTGAAGAACAG 842  
Qy 842 GCCTTGGGCAACTCTGTTTAAAGACGTTTGGCAGGGGCTACGCGCGCTCTGCTCGAT 901  
Db 843 GCCTTGGGCAACTCTGTTTAAAGACGTTTGGCAGGGGCTACGCGCGCTCTGCTCGAT 902  
Qy 902 CAGAACTCCACAAACTGATCAGCCCCCTTGCATGTCTGAACACAGTGTGGAACCTGCACC 961  
Db 903 CAGAACTCCACAAACTGATCAGCCCCCTTGCATGTCTGAACACAGTGTGGAACCTGCACC 962  
Qy 962 ACCCCGAGGACGAGGCCCCCTGCCCCCTGCCCCAGCAGCCCTTCCCTATAGCAGACTGC 1021  
Db 963 ACCCCGAGGACGAGGCCCCCTGCCCCCTGCCCCAGCAGCCCTTCCCTATAGCAGACTGC 1022  
Qy 1022 CTGATCCCTTCCCATTCACACCTCTCCAGCCCTTGAAACCTTGAAACCTCAGCCCTGAGTCCCTTC 1081  
Db 1023 CTGATCCCTTCCCATTCACACCTCTCCAGCCCTTGAAACCTTGAAACCTCAGCCCTGAGTCCCTTC 1082  
Qy 1082 TGGGCAAACTGGCCCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACCTGAGCAAAAC 1141  
Db 1083 TGGGCAAACTGGCCCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACCTGAGCAAAAC 1142  
Qy 1142 TGGCCTGTGTAGACAGTCAAAGCCCTGCTGGCCCAACCTGGAGCCCAAGCTGCCTGT 1201  
Db 1143 TGGCCTGTGTAGACAGTCAAAGCCCTGCTGGCCCAACCTGGAGCCCAAGCTGCCTGT 1202  
Qy 1202 CTGCTGTGTCATGAGAGTTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGCA 1261  
Db 1203 CTGCTGTGTCATGAGAGTTTCTGTGGAGGATACCTAGTGCATGCTCTGCAAGGCA 1262  
Qy 1262 GCCTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGCCCAAGCCTGGGACAGCGGG 1321  
Db 1263 GCCTGAGCTCAAGCCAGGCCACAGCCTGACAGCCTGGCCCAAGCCTGGGACAGCGGG 1322  
Qy 1322 GCTCCAGATCCCGGAGCCAGCCCAAAACTGAGGACAAAGAGGCTGCTGCTCACTG 1381  
Db 1323 GCTCCAGATCCCGGAGCCAGCCCAAAACTGAGGACAAAGAGGCTGCTGCTCACTG 1382  
Qy 1382 AGAACTCAAGCCAGTGTATGATGATACCGAGAGAGTCCATGGGCCAGCAGCAGC 1441  
Db 1383 AGAACTCAAGCCAGTGTATGATGATACCGAGAGAGTCCATGGGCCAGCAGCAGC 1442  
Qy 1442 TCCGCTGGGACAGAGCTCTTCGAGAGGTGACAGAGTGGAGGACAGCAGACTGGCT 1501  
Db 1443 TCCGCTGGGACAGAGCTCTTCGAGAGGTGACAGAGTGGAGGACAGCAGACTGGCT 1502  
Qy 1502 TCCAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCAT 1561  
Db 1503 TCCAGTGGCTGTCAAAGAGTGGCCCTGGAAGTATTTTCGGGACAGAGAGCTGATGGCAT 1562  
Qy 1562 GTGACAGATTGACCTCACCGAGAAATTTGCTCCCTTGTATGAGAGTGTGAGAGAGGGCTT 1621  
Db 1563 GTGACAGATTGACCTCACCGAGAAATTTGCTCCCTTGTATGAGAGTGTGAGAGAGGGCTT 1622  
Qy 1622 GGGTCAACATCTTCATGAGCTGTGGAAGGTGCTCCCTGGGCGAGCTGTGCAAGGAGC 1681  
Db 1623 GGGTCAACATCTTCATGAGCTGTGGAAGGTGCTCCCTGGGCGAGCTGTGCAAGGAGC 1682  
Qy 1682 AGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTTGGGCGAGGCCCTGGAGGGTCTGG 1741  
Db 1683 AGGGCTGTCTCCAGAGACCGGGCCCTGTACTACTTGGGCGAGGCCCTGGAGGGTCTGG 1742  
Qy 1742 AATACCTCACTCAAGAGGATTTCTGATGAGGAGCTGCAAGAGCTGCAAGCTGCTCTGT 1801  
Db 1743 AATACCTCACTCAAGAGGATTTCTGATGAGGAGCTGCAAGAGCTGCAAGCTGCTCTGT 1802

Qy 1802 CCAGCGATGGAGGCCACGACAGCCCTCTGTGATCTTTTGGCCATGCTGTGTCTTCAACCTG 1861  
Db 1803 CCAGCGATGGAGGCCACGACAGCCCTCTGTGATCTTTTGGCCATGCTGTGTCTTCAACCTG 1862  
Qy 1862 ATGSCCTTGGGAAAGTCTCTCTCACAGGGAGCTACATCCCTGGCCACAGAGACCCACATGG 1921  
Db 1863 ATGSCCTTGGGAAAGTCTCTCTCACAGGGAGCTACATCCCTGGCCACAGAGACCCACATGG 1922  
Qy 1922 CTCCGAGGTGGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGAGAGAGCTGCT 1981  
Db 1923 CTCCGAGGTGGTCTGGGAGGAGCTGCGACGCAAGGTGGATGTCTGAGAGAGCTGCT 1982  
Qy 1982 GTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTCGAGGGGCGC 2041  
Db 1983 GTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTTTCGAGGGGCGC 2042  
Qy 2042 TCTGCTCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCACTCTCTGCGGCC 2101  
Db 2043 TCTGCTCTCAAGATTGCCAGGAGCTCCGCTGTGAGGGAGATCCCACTCTCTGCGGCC 2102  
Qy 2102 CTCTCAGAGCCAGGCCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTG 2161  
Db 2103 CTCTCAGAGCCAGGCCCATCCAGAGGGGCTGAGGAAAGAGCCCATCCACCGCTGTCTG 2162  
Qy 2162 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTCTGAAGAGCC 2221  
Db 2163 CAGCGAGCTGGGAGGAGGTGAACCGGGCACTTACAGCAAGTGGGAGGTCTGAAGAGCC 2222  
Qy 2222 CTTGGAGGGAGATATATAAGAACCAAGACATCCACCGCCCAATCAAGCCCAATTTACCAC 2281  
Db 2223 CTTGGAGGGAGATATATAAGAACCAAGACATCCACCGCCCAATCAAGCCCAATTTACCAC 2282  
Qy 2282 AGACCTTCCATGCCCAGCAGAGAGCTTTTCGCAAGGGCCCGAGGCCCCGGCCAGCTG 2341  
Db 2283 AGACCTTCCATGCCCAGCAGAGAGCTTTTCGCAAGGGCCCGAGGCCCCGGCCAGCTG 2342  
Qy 2342 AGGAGACAACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCCAAGAGAGCCCCAGAGC 2401  
Db 2343 AGGAGACAACAGGAGAGCCCTTAAGCTCCAGCCCTCTCTCCCAAGAGAGCCCCAGAGC 2402  
Qy 2402 CAACCAAGTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGATGTGGGAAACCTTTAC 2461  
Db 2403 CAACCAAGTCTCTCTCTTGAATTTGAGCAAGGAGGAGTCTGGATGTGGGAAACCTTTAC 2462  
Qy 2462 CTCTGCTCTCTCTGGAGCAGCCCTCCAGAAACCCAGCTCACAGAGCGGAAAGCAA 2521  
Db 2463 CTCTGCTCTCTCTGGAGCAGCCCTCCAGAAACCCAGCTCACAGAGCGGAAAGCAA 2522  
Qy 2522 CCGTCCCGAGAGCAGGAACTCAGCAGCTGGAAATAGAAATTTCTCTCAACAGCCTGTCCC 2581  
Db 2523 CCGTCCCGAGAGCAGGAACTCAGCAGCTGGAAATAGAAATTTCTCTCAACAGCCTGTCCC 2582  
Qy 2582 AGCCATTTTCTTGGAGGAGCGAGCAAAATTTCTCTGCTGCTCAGCATCCAGAGCCTCT 2641  
Db 2583 AGCCATTTTCTTGGAGGAGCGAGCAAAATTTCTCTGCTGCTCAGCATCCAGAGCCTCT 2642  
Qy 2642 CCCTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCCGCGGACACCC 2701  
Db 2643 CCCTGTCCGATGACAGTGAAGAACCCATCAAGGCTCTCAAGCTCCGCGGACACCC 2702  
Qy 2702 TGAGCTCAGCGCTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGAACA 2761  
Db 2703 TGAGCTCAGCGCTACACTCTCTGGAGCAGCCAGGCGGAGGCTCGAAGCTCCAGCTGGAACA 2762  
Qy 2762 TGCTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCC 2821  
Db 2763 TGCTGTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCAATGTGTGAAAGTCC 2822  
Qy 2822 AATACAGTCTCTTAATGTGAAACCTTCACATCCCGGAGTTCACCGGCTCAAGTGG 2881  
Db 2823 AATACAGTCTCTTAATGTGAAACCTTCACATCCCGGAGTTCACCGGCTCAAGTGG 2882



Db 53 CCGGGGACTGTGCGGTGTGGAACTGTAGCTGTGTAGAGGTGGACTCTGTATCAATTGA 112  
Qy 62 GGATGTTTGGAGGATAGTATGTGTGAGAGGACACATAAAACAGGACAGACCTTTTG 121  
Db 113 GGATGTTTGGAGGATAGTATGTGTGAGAGGACACATAAAACAGGACAGACCTTTTG 172  
Qy 122 CCCCTGCTTTCTCCCAACCCCAAGGTGACCTGTGTTCTCCAGGTCTGGGATTTCTAA 181  
Db 173 CCCCTGCTTTCTCCCAACCCCAAGGTGACCTGTGTTCTCCAGGTCTGGGATTTCTAA 232  
Qy 182 GTGACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGACAAAGCTGGGAGATGGCAGTGA 241  
Db 233 GTGACCTGCTCTGTGTTTGGTCTCTCTCAGGATGAGACAAAGCTGGGAGATGGCAGTGA 292  
Qy 242 TGGAAATGGCTGCCAGGTGCCCTTGCTCAGCAGTGGGGCAGCAGAGAACTCCCA 301  
Db 293 TGGAAATGGCTGCCAGGTGCCCTTGCTCAGCAGTGGGGCAGCAGAGAACTCCCA 352  
Qy 302 AGCCAAAGGAGAACCGCCGACCTGGGGAGAAACAGAGTCCGTCTACAGCTTGAGG 361  
Db 353 AAGCCAAAGGAGAACCGCCGACCTGGGGAGAAACAGAGTCCGTCTACAGCTTGAGG 412  
Qy 362 CCGTGGAGAACAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACCA 421  
Db 413 CCGTGGAGAACAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACCA 472  
Qy 422 AGGACACAGCCAAAGGAGCTCCGAGGAGGCGGAGCTGCCATCTCTATCATGCCCAGG 481  
Db 473 AGGACACAGCCAAAGGAGCTCCGAGGAGGCGGAGCTGCCATCTCTATCATGCCCAGG 532  
Qy 482 CTGAGTGTGAATAGCAAGAGTTTCAAGCCCACTTTTCAAGCGATTTTCAATGCGTG 541  
Db 533 CTGAGTGTGAATAGCAAGAGTTTCAAGCCCACTTTTCAAGCGATTTTCAATGCGTG 592  
Qy 542 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGCCCATG 601  
Db 593 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGCCCATG 652  
Qy 602 CTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAGGGAAAGCGTCGACAGCAAGCCCGGA 661  
Db 653 CTACAGAGGGCAAAATGGCCCGTGTGTGTTGGAGGGAAAGCGTCGACAGCAAGCCCGGA 712  
Qy 662 AGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGAGAGAGTGGCTTGGCCAAAC 721  
Db 713 AGAAACGGAAGAAAGAGCTCAAAGTCCCTGGCTCATGAGAGAGTGGCTTGGCCAAAC 772  
Qy 722 CCCTCCCAAGACCCCTGAGCAGAGAGCTGCACCATCCAGTCCAGAGGATGAGTCTC 781  
Db 773 CCCTCCCAAGACCCCTGAGCAGAGAGCTGCACCATCCAGTCCAGAGGATGAGTCTC 832  
Qy 782 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTGAAGAACAG 841  
Db 833 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAAGCTCTGAAGAACAG 892  
Qy 842 GCCTTGGGCAACTCTGTTTAAAGCAGTGTGGCAGGCGCTACAGCCCGGCTCTGCTCGAT 901  
Db 893 GCCTTGGGCAACTCTGTTTAAAGCAGTGTGGCAGGCGCTACAGCCCGGCTCTGCTCGAT 952  
Qy 902 CAGAACTCCACAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGAAACTGCAAC 961  
Db 953 CAGAACTCCACAACTGATCAGCCCTTGCAATGTCTGAACCAAGTGTGAAACTGCAAC 1012  
Qy 962 ACCCCAGGAGGAGGCGCCCTGCTCCCAAGCAGCAGCCCTTCCCTATAGCAGACTGC 1021  
Db 1013 ACCCCAGGAGGAGGCGCCCTGCTCCCAAGCAGCAGCCCTTCCCTATAGCAGACTGC 1072  
Qy 1022 CTCAATCCCTTCCCAATTCACCCCTCTCCAGCCCTTGGAAACCTCACCTCTGAGTCTTCC 1081  
Db 1073 CTCAATCCCTTCCCAATTCACCCCTCTCCAGCCCTTGGAAACCTCACCTCTGAGTCTTCC 1132  
Qy 1082 TGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGCCTGACCCACACTGAGCAAC 1141  
Db 1133 TGGGCAAACTGGCCTGTGTAGACAGCCAGAAACCCCTTGCCTGACCCACACTGAGCAAC 1192

Qy 1142 TGGCCTGTGTAGACAGTCCAAAGCCCTTGCTGCTGCCCCACACCTGGAGCCACAGTGCCTGT 1201  
Db 1193 TGGCCTGTGTAGACAGTCCAAAGCCCTTGCTGCTGCCCCACACCTGGAGCCACAGTGCCTGT 1252  
Qy 1202 CTGCTGTGCTCCCATGAAAGTCTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1261  
Db 1253 CTGCTGTGCTCCCATGAAAGTCTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1312  
Qy 1262 GCCTGAGCTCAAGCAGGCGCCACAGACCTGACCCAGCTGGCCCAAGACCTGGGSCAGCAGGG 1321  
Db 1313 GCCTGAGCTCAGGCCAGGCCACAGACCTGACCCAGCTGGCCCAAGACCTGGGSCAGCAGGG 1372  
Qy 1322 GTTCAGATCCCGGAGGCCACAGCCGCCAAATCTGAGGACAAAGAGGTGTCTGCTCACTG 1381  
Db 1373 GCTCCAGATCCCGGAGGCCACAGCCGCCAAATCTGAGGACAAAGAGGTGTCTGCTCACTG 1432  
Qy 1382 AGAACTCAAGCAGTGGATTATGATACCGAGGAAGTCCACTGGGSCACGACCCAGC 1441  
Db 1433 AGAACTCAAGCAGTGGATTATGATACCGAGGAAGTCCACTGGGSCACGACCCAGC 1492  
Qy 1442 TCCGCTGGGACAGAGGCTCCTTCCGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCT 1501  
Db 1493 TCCGCTGGGACAGAGGCTCCTTCCGAGAGGTGCACAGGATGGAGGACAAAGCAGACTGGCT 1552  
Qy 1502 TCCAGTGCCTGTCAAAGAGGTGGCTGGAAGTATTTCCGGGACAGAGAGCTGATGGCAT 1561  
Db 1553 TCCAGTGCCTGTCAAAGAGGTGGCTGGAAGTATTTCCGGGACAGAGAGCTGATGGCAT 1612  
Qy 1562 GTGACAGATTGACCTCACCCAGAAATGTCTCCTTTGATGAGAGTGTGAGAGAGGGCTTT 1621  
Db 1613 GTGACAGATTGACCTCACCCAGAAATGTCTCCTTTGATGAGAGTGTGAGAGAGGGCTTT 1672  
Qy 1622 GGGTCAACATCTTTCATGGAGCTGTGGAAGTGTCTTCCCTGGGACAGCTGTGCAAGGAGC 1681  
Db 1673 GGGTCAACATCTTTCATGGAGCTGTGGAAGTGTCTTCCCTGGGACAGCTGTGCAAGGAGC 1732  
Qy 1682 AGGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGACAGGCGCTGGAGGGTCTGG 1741  
Db 1733 AGGCTGTCTCCAGAGGACCGGCGCTGTACTACCTGGGACAGGCGCTGGAGGGTCTGG 1792  
Qy 1742 AATACCTCACTCAGAGAGATTCTGCATGGGAGCTCAAGAGCTGACAACTGCTCTCTGT 1801  
Db 1793 AATACCTCACTCAGAGAGATTCTGCATGGGAGCTGCAAGAGCTGCAACGCTGCTCTGT 1852  
Qy 1802 CCAGCATGGGAGCCAGCAGCTCTGTCACCTTGGCCATGCTGTGCTCTTCAACCTG 1861  
Db 1853 CCAGCATGGGAGCCAGCAGCTCTGTCACCTTGGCCATGCTGTGCTCTTCAACCTG 1912  
Qy 1862 ATGGCTTGGGAAAGTCTTCTCAGAGGAGTACATCCCTGGCAGACAGACCCACATGG 1921  
Db 1913 ATGGCTTGGGAAAGTCTTCTCAGAGGAGTACATCCCTGGCAGACAGACCCACATGG 1972  
Qy 1922 CTCGAGAGTGTGCTGGGACAGAGCTGCACGCAAGGTGGATGTCTGAGCAGCTGCT 1981  
Db 1973 CTCGAGAGTGTGCTGGGACAGAGCTGCACGCAAGGTGGATGTCTGAGCAGCTGCT 2032  
Qy 1982 GTATGATGTCACATGCTCAACGGCTGCCACCCCTGGAATCAGTTCTTCCGAGGGCGC 2041  
Db 2033 GTATGATGTCACATGCTCAACGGCTGCCACCCCTGGAATCAGTTCTTCCGAGGGCGC 2092  
Qy 2042 TCTGCTTCAAGATTGCGCAGGAGCTCCGCTGTGAGGAGATCCCACTCTTCTGCGGCC 2101  
Db 2093 TCTGCTTCAAGATTGCGCAGGAGCTCCGCTGTGAGGAGATCCCACTCTTCTGCGGCC 2152  
Qy 2102 CTCTCAGCCCGGAGCCATCCAAAGAGGGTGTGAGGAAAGAGCCCATCAACGCTGTGCTG 2161  
Db 2153 CTCTCAGCCCGGAGCCATCCAAAGAGGGTGTGAGGAAAGAGCCCATCAACGCTGTGCTG 2212  
Qy 2162 CAGGAGCTGGGAGGAGGTGAAACGGGCACTACAGCAAGTGGAGGCTGTGAGAGCC 2221  
Db 2213 CAGGAGCTGGGAGGAGGTGAAACGGGCACTACAGCAAGTGGAGGCTGTGAGAGCC 2272

QY 2222 CTTGGAGGGGAGATATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACC 2281  
DB 2273 CTTGGAGGGGAGATATATAAGAAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACC 2332  
QY 2282 AGACCCCTCATGCGCCAGCGAGAGCTTTTCGCCAAGGCGCCCAAGGCGCCCGCGCCAGCTGG 2341  
DB 2333 AGACCCCTCATGCGCCAGCGAGAGCTTTTCGCCAAGGCGCCCAAGGCGCCCGCGCCAGCTGG 2392  
QY 2342 AGGAGACACAGCGAGCGCCCTTAGCTCCAGCTCCCTCCACCCAGAGCGCCCGCCAGAGC 2401  
DB 2393 AGGAGACACAGCGAGCGCCCTTAGCTCCAGCTCCCTCCACCCAGAGCGCCCGCCAGAGC 2452  
QY 2402 CAAAACAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTTTAC 2461  
DB 2453 CAAAACAAGTCTCTCCCTTGAATTTGAGCAAGGAGAGTCTGGGATGTGGGAACCTTTAC 2512  
QY 2462 CTCTGTCTCTCTGGAGCGAGCGCCCTTGCCAGAAAACCCAGCTCAACAGAGCGGAAAGCAA 2521  
DB 2513 CTCTGTCTCTCTGGAGCGAGCGCCCTTGCCAGAAAACCCAGCTCAACAGAGCGGAAAGCAA 2572  
QY 2522 CGGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTTCTCAACAGCCTGTCCC 2581  
DB 2573 CGGTCCCGGAGCAGGAACCTGCAGCAGCTGGAAATAGAAATTTCTCAACAGCCTGTCCC 2632  
QY 2582 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCCTCT 2641  
DB 2633 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATTTCTCTGTGCTCAGCATCGACAGCCTCT 2692  
QY 2642 CCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGCCCTCTCAAAAGCTCGCGGACACCC 2701  
DB 2693 CCCTGTGCGATGACAGTGAGAGAAACCCATCAAAAGCCCTCTCAAAAGCTCGCGGACACCC 2752  
QY 2702 TGAGCTCAGGCGGTACACTCTCTGGAGCAGCGAGCGGAGCTCGAAGCTCCAGCTGGAACA 2761  
DB 2753 TGAGCTCAGGCGGTACACTCTCTGGAGCAGCGAGCGGAGCTCGAAGCTCCAGCTGGAACA 2812  
QY 2762 TGGTGTGCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGAAAGTCC 2821  
DB 2813 TGGTGTGCGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATATGTTGAAAGTCC 2872  
QY 2822 AAATACAGTCTCTTAATGTTGAACACCTTGCAATCCGGAGTTCACCGGGTCAAAGTGG 2881  
DB 2873 AAATACAGTCTCTTAATGTTGAACACCTTGCAATCCGGAGTTCACCGGGTCAAAGTGG 2932  
QY 2882 GAGACATCGCCACTGSCATCAGCAGCGAGATCCAGCTCGAGCTTCAGCTTGTGCACCA 2941  
DB 2933 GAGACATCGCCACTGSCATCAGCAGCGAGATCCAGCTCGAGCTTCAGCTTGTGCACCA 2992  
QY 2942 AAGACGGGCGAGCCTGTTCTGTACGACATGGAGGTGCCAGACTCGGGCATTCGACCTGCAAT 3001  
DB 2993 AAGACGGGCGAGCCTGTTCTGTACGACATGGAGGTGCCAGACTCGGGCATTCGACCTGCAAT 3052  
QY 3002 GCACACTGCGCCCTGATGSCAGCTTCGCTGAGCTGGAGGTCAAGCATGCGCCAGCTGG 3061  
DB 3053 GCACACTGCGCCCTGATGSCAGCTTCGCTGAGCTGGAGGTCAAGCATGCGCCAGCTGG 3112  
QY 3062 AGAAAGGCGCCCTAAACCTGCTCCACCGCCGCTCCACACTGCGCGGAAAGCAGCCTTCC 3121  
DB 3113 AGAAAGGCGCCCTAAACCTGCTCCACCGCCGCTCCACACTGCGCGGAAAGCAGCCTTCC 3171  
QY 3122 TGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCCGTTCAGGGGATTCGCCAG 3181  
DB 3172 TGCTCGGTGACGATGCTGCCCTGAAACACAGGCTCAGCCGTTCAGGGGATTCGCCAG 3230  
QY 3182 CCCCCCGGTCAAGTGGGAAACAGGGCTTCGACGAGCAAGGTGGGGGCAAGCAGAATG 3241  
DB 3231 -CCCCCGGTCAAGTGGGAAACAGGGCTTCGACGAGCAAGGTGGGGGCAAGCAGAATG 3289  
QY 3242 CCTCCCGAGATTTCACTGAGCGCTGCCCGCCAGCTCTGCTGAAACAAAC - TCCGCGACGT 3300  
DB 3290 CCTCCCGAGATTTCACTGAGCGCTGCCCGCCAGCTCTGCTGAGAAACAACTCCGCGACGT 3349  
QY 3301 GAAGAGACAGAGGAGGATGGCAGGAGTTTACCTGGGGGAAACAAACACAGGGATCTTTTCT 3360

DB 3350 GAAGAGACAG - AGGAGGATGGCAGGATTAACCTCGGGAACAAACAGG - - - - ATCTTCTC 3404  
QY 3361 GCCCCTGTCTCAGTCGAGTTGGCTGACCCCGTTGGATCAGTGACCAATTTGTTGGCAGAC 3420  
DB 3405 TGCCCTGTCTCCAGTCAGTTGGCTGACCCCGTTGGATCAGTGACCAATTTGTTGGCAGAC 3464  
QY 3421 AGGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGGGCCCTCACC 3480  
DB 3465 AGGGGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGCGAGCCCTTCGGGCCCTCACC 3524  
QY 3481 CTCACAGGCTGTCTGAGAGTGTCAAGTGTGTAAAGGCGCCAAACTCAGGTTTCAGTCAGAA 3540  
DB 3525 CTCACAGGCTGTCTGAGAGTGTCAAGTGTGTAAAGGCGCCAAACTCAGGTTTCAGTCAGAA 3584  
QY 3541 CCAGTGTACAGAGGTATGCGCCCGCTAGGTTAAAGGGGGCCCTCTAAAACCCCTTGGCTGGC 3600  
DB 3585 CCAGTGTACAGAGGTATGCGCCCGCTAGGTTAAAGGGGGCCCTCTAAAACCCCTTGGCTGGC 3644  
QY 3601 CTCACCTGCGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAG 3660  
DB 3645 CTCACCTGCGCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTGACCCCTGGGAG 3704  
QY 3661 GCTCCCTGTGTAGAATACACCACTTTTCAGGTTGTTCACACACAGGTCCTGAGTTGACC 3720  
DB 3705 GCTCCCTGTGTAGAATACACCACTTTTCAGGTTGTTCACACACAGGTCCTGAGTTGACC 3764  
QY 3721 TCTGTTTCAGCCAGGACCAAGAGGTGTGTAGTGAAGTGTCTCAGTCCCCAGACA 3780  
DB 3765 TCTGTTTCAGCCAGGACCAAGAGGTGTGTAGTGAAGTGTCTCAGTCCCCAGACA 3824  
QY 3781 TGTGCCCCCTTTGCTGCTGCTACCACTCTCCAGAGCAGCAGCCCGGAGCCCTTCA 3840  
DB 3825 TGTGCCCCCTTTGCTGCTGCTACCACTCTCTCCAGAGCAGCAGCCCGGAGCCCTTCA 3884  
QY 3841 GGCCAGCAGCTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGTGAAGGT 3900  
DB 3885 GGCCAGCAGCTGCCCCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGTGAAGGT 3944  
QY 3901 GATGAGGATATGCTGTGACAGAAACAGTCTGTGAGTGGACATGATCAGTGTGAAGAAAG 3960  
DB 3945 GATGAGGATATGCTGTGACAGAAACAGTCTGTGAGTGGACATGATCAGTGTGAAGAAAG 4004  
QY 3961 CAGCAGAGAGAGAG - TCCGGCGCCCGAGCCCACTATCAGTGTCCAGCTGCTGCTGCC 4019  
DB 4005 CAGCAGAGAGAGAGCCTCCGGCGCCCGAGCCCACTATCAGTGTCCAGGCTGCTGCTGCC 4064  
QY 4020 CCAGAGCAGCTCAGCATCACTGACACTCACCTTCCCTGCCCCCTGGCCAGAGGTA 4079  
DB 4065 CCAGAGCAGCTCAGCATCACTGACACTCACCTTCCCTGCCCCCTGGCCAGAGGTA 4124  
QY 4080 GTGCCGACGGCACTTTTGCACTCTGATGACCTCAAGACATTTTCATGGCTGCCCTTGGA 4139  
DB 4125 GTGCCGACGGCACTTTTGCACTCTGATGACCTCAAGACATTTTCATGGCTGCCCTTGGA 4184  
QY 4140 GGGCAGGCGCAGGAGTGCACCTGTAGGAGCATAGCAGCAGGAGATGGGGTGAAGGA 4199  
DB 4185 GGGCAGGCGCAGGAGTGCACCTGTAGGAGCATAGCAGCAGGAGATGGGGTGAAGGA 4244  
QY 4200 CACAGCTTTGAGCTGTCCACATGATGTGACTCTCTCAAAACCTCTTCCAGATTTCTTAAG 4259  
DB 4245 CACAGCTTTGAGCTGTCCACATGATGTGACTCTCTCAAAACCTCTTCCAGATTTCTTAAG 4304  
QY 4260 AATAGCAGCCCTTCCCATTTGCCCACTAGCCCTTCTCCAGGGGAGTACTCAGG 4319  
DB 4305 AATAGCAGCCCTTCCCATTTGCCCACTAGCCCTTCTCCAGGGGAGTACTCAGG 4364  
QY 4320 ACTCAGTAGGATTAATCAGCTGTGAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCC 4379  
DB 4365 ACTCAGTAGGATTAATCAGCTGTGAATCGTCAGGGGTGTCTGCTAGCCTCAACCTCC 4424  
QY 4380 TGGGCGAGGAGCGCCGAGACTCCGTGGGAGAGCTCAATCCACATCTTGCACAGACAG 4439



Db 4425 TGGGCGAGGAGCGCGAGACTCCGTGGGAGAGCTCATTTCCACATCTTCCCAAGACAG 4484  
Qy 4440 CCTTTGTCCAGCTGTGCACATTTAGTTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCCA 4499  
Db 4485 CCTTTGTCCAGCTGTGCACATTTAGTTCAGACTGCTCCCGGGAGAGAGCCCGGCCCCCA 4544  
Qy 4500 GCACATAAGAACTGCGAGCCTTGTGACTGACAGTCTGGGTGTAGAGAACTCTTTGTAA 4559  
Db 4545 GCACATAAGAACTGCGAGCCTTGTGACTGACAGTCTGGGTGTAGAGAACTCTTTGTAA 4604  
Qy 4560 GCAATAAAGTTTGGGTGATGACAAATGTT 4589  
Db 4605 GCAATAAAGTTTGGGTGATGACAAATGTT 4634

RESULT 15  
US-10-170-235-9152  
; Sequence 9152 Application US/10170235  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig  
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN  
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF  
; FILE REFERENCE: CL001380  
; CURRENT APPLICATION NUMBER: US/10/170,235  
; CURRENT FILING DATE: 2003-03-17  
; NUMBER OF SEQ ID NOS: 42514  
; SEQ ID NO 9152  
; LENGTH: 4684  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-10-170-235-9152

Query Match 96.6%; Score 4440; DB 49; Length 4684;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 4549; Conservative 0; Mismatches 30; Indels 11; Gaps 8;

Qy 3 CGGGGGAGCTGTGCGGTGTGAAAGTGTAGTGTGA-AGGTGAGTCTGTGTTACCATGA 61  
Db 53 CGGGGGAGCTGTGCGGTGTGAAAGTGTAGTGTGAAGTGTGAGAGTGTGACTCTGTGTTACCATGA 112  
Qy 62 GGATGTTTGGAGGATGATGTGTGGCAGAGGACACATATAACAGGACAGAGCCCTTTG 121  
Db 113 GGATGTTTGGAGGATGATGTGTGGCAGAGGACACATATAACAGGACAGAGCCCTTTG 172  
Qy 122 CCCCTGCTTCTCCCAAGGCTGACCTGTGTCTCCAGAGTCTGGGATTTCTAA 181  
Db 173 CCCCTGCTTCTCCCAAGGCTGACCTGTGTCTCCAGAGTCTGGGATTTCTAA 232  
Qy 182 GTGACCTGCTGTGTGTGTGTCTCTCAGGATGAGCAAGCTGGGAGATGGCAGTGA 241  
Db 233 GTGACCTGCTGTGTGTGTGTCTCTCAGGATGAGCAAGCTGGGAGATGGCAGTGA 292  
Qy 242 TGGAAATGGCTGCGCAGGTGCGCTGCTCAGAGTGGGGCAGCAGAGAACTCCCA 301  
Db 293 TGGAAATGGCTGCGCAGGTGCGCTGCTCAGAGTGGGGCAGCAGAGAACTCCCA 352  
Qy 302 AGCCAAAGGAGAGAGCGCGCACTGGGGAAGAAACAGAGTCTGCTCTACAAGCTTCAGG 361  
Db 353 AAGCCAAAGGAGAGAGCGCGCACTGGGGAAGAAACAGAGTCTGCTCTACAAGCTTCAGG 412  
Qy 362 CCGTGGAGAGAGCGCTGTGTGTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACCA 421  
Db 413 CCGTGGAGAGAGCGCTGTGTGTCTGCGGAAAGTGGGAGATCCTGAATGACGTGATTACCA 472  
Qy 422 AGGCGACAGCGAAGAGGCTCCAGGAGGCGCAGTGCATCTCTATCATGCCCCAGG 481  
Db 473 AGGCGACAGCGAAGAGGCTCCAGGAGGCGCAGTGCATCTCTATCATGCCCCAGG 532  
Qy 482 CTGAGTGTGAATAGCCAGAGTTCAAGCCCTTTTCAGAGCGATTTTCATCGCTG 541  
Db 533 CTGAGTGTGAATAGCCAGAGTTCAAGCCCTTTTCAGAGCGATTTTCATCGCTG 592  
Qy 542 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGTATCAGATCCCCCAACAAATGTGGCCCATG 601

Db 593 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCCCAACAAATGTGCCCATG 652  
Qy 602 CTACAGAGGCGAAATATGGCCCGTGTGTGTGGAGGAAAGCGTCGACGAAACCCCGGA 661  
Db 653 CTACAGAGGCGAAATATGGCCCGTGTGTGTGGAGGAAAGCGTCGACGAAACCCCGGA 712  
Qy 662 AGAAACGGAAGAAAGAGAGTCAAAAGTCCCTGTGCTCATGAGAGTGGCTTGGCCAAAC 721  
Db 713 AGAAACGGAAGAAAGAGAGTCAAAAGTCCCTGTGCTCATGAGAGTGGCTTGGCCAAAC 772  
Qy 722 CCCTCCCGCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGTCAAGTCTC 781  
Db 773 CCCTCCCGCAGGACCCCTGAGCAGAGAGTGCACCATCCAGTGCAGAGAGTCAAGTCTC 832  
Qy 782 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTCAAGAAACAG 841  
Db 833 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTCACCAAGCTCTCAAGAAACAG 892  
Qy 842 GCCTTGGGCAACTCTGTTTTAAGCAGCTTGGGAGGCGCTACGCGCGCTCTGCTCGAT 901  
Db 893 GCCTTGGGCAACTCTGTTTTAAGCAGCTTGGGAGGCGCTACGCGCGCTCTGCTCGAT 952  
Qy 902 CAGAACTCCAAACACTGATCAGCCCTTGAATGTCTGAACCACTGTGGAACCTGCACC 961  
Db 953 CAGAACTCCAAACACTGATCAGCCCTTGAATGTCTGAACCACTGTGGAACCTGCACC 1012  
Qy 962 ACCCCAGAGAGGAGGCGCCCTGCGCTGCGCAGCAGCAGCCCTTCCCTATAGCAGACTGC 1021  
Db 1013 ACCCCAGAGAGGAGGCGCCCTGCGCTGCGCAGCAGCAGCCCTTCCCTATAGCAGACTGC 1072  
Qy 1022 CTATCCCTTCCATTCACCCCTCTCCAGCCCTGGAAACCTCACCTCTGTGAGTCTCTCC 1081  
Db 1073 CTATCCCTTCCATTCACCCCTCTCCAGCCCTGGAAACCTCACCTCTGTGAGTCTCTCC 1132  
Qy 1082 TGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACACCTGAGCAAC 1141  
Db 1133 TGGGCAAACTGGCTGTGTAGACAGCCAGAAACCTTGGCTGACCCACACCTGAGCAAC 1192  
Qy 1142 TGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGT 1201  
Db 1193 TGGCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGT 1252  
Qy 1202 CTGTGTGCTCCATGAGAAAGTCTTGTGGAGGAATACCTAGTGCATGCTCTGCAAGCA 1261  
Db 1253 CTGTGTGCTCCATGAGAAAGTCTTGTGGAGGAATACCTAGTGCATGCTCTGCAAGCA 1312  
Qy 1262 GCGTGTGCTCAAGCCAGGCGCCACAGCTGACAGCCCTGGCCCAAGACCTGGGAGCAAGG 1321  
Db 1313 GCGTGTGCTCAAGCCAGGCGCCACAGCTGACAGCCCTGGCCCAAGACCTGGGAGCAAGG 1372  
Qy 1322 GCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAACTGAGGGTGTCTGCTCACTG 1381  
Db 1373 GCTCCAGATCCCGGAGCCAGCCCAAACTGAGGACAACTGAGGGTGTCTGCTCACTG 1432  
Qy 1382 AGAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCCACGACCAAGC 1441  
Db 1433 AGAACTCAAGCCAGTGGATTTATGAGTACCGAGAAAGTCCACTGGGCCACGACCAAGC 1492  
Qy 1442 TCCGCTTGGGAGAGGCTCTTGGAGAGTGCACAGGATGGAGGACAGCAGACTGGCT 1501  
Db 1493 TCCGCTTGGGAGAGGCTCTTGGAGAGTGCACAGGATGGAGGACAGCAGACTGGCT 1552  
Qy 1502 TCCAGTGGCTGTCAAAAGGTCGCTGGAAGTATTTTCGGGAGAGGAGTGTATGAT 1561  
Db 1553 TCCAGTGGCTGTCAAAAGGTCGCTGGAAGTATTTTCGGGAGAGGAGTGTATGAT 1612  
Qy 1562 GTGAGGATTTGACCTCACCCAGAAATTTCTTGTATGAGTGTGTGAGAAAGGCTT 1621  
Db 1613 GTGAGGATTTGACCTCACCCAGAAATTTCTTGTATGAGTGTGTGAGAAAGGCTT 1672  
Qy 1622 GGGTCAACATCTTATGAGAGTGTGTGAAGGTGCTCCCTGGGCGAGCTGTGTGAAGAGC 1681

Db	1673	GGGTCAACATCTTCATGGAGCTGCTGGAAGGTGGCTCCCTGTGGGCCACGCTGGTCAAGAGC	1732
Qy	1682	AGGGCTGTCTCCAGAGAACCGGGCCCTGTGTACTACCTGGGCCAGGCCCTCGAGGGTCTGG	1741
Db	1733	AGGGCTGTCTCCAGAGAACCGGGCCCTGTGTACTACCTGGGCCAGGCCCTCGAGGGTCTGG	1792
Qy	1742	AATACCTTCCATCAAGAAAGATTCGTGATGGGGACGTCAAAGCTGACAAACGTGCTCTGT	1801
Db	1793	AATACCTTCCATCAAGAAAGATTCGTGATGGGGACGTCAAAGCTGACAAACGTGCTCTGT	1852
Qy	1802	CCAGCATGGGAGCACCGACCCCTCTGTGACTTTTGGCCATGCTGTGTGTCTTCAACCTG	1861
Db	1853	CCAGCATGGGAGCACCGACCCCTCTGTGACTTTTGGCCATGCTGTGTGTCTTCAACCTG	1912
Qy	1862	ATGGCTTGGGAAAGTCTCTTGCTCACAGGGGACTACATCCCTGGCACAGACACCAATGG	1921
Db	1913	ATGGCTTGGGAAAGTCTCTTGCTCACAGGGGACTACATCCCTGGCACAGACACCAATGG	1972
Qy	1922	CTCCGAGGTGGTCTGTGGCAGGAGCTTGCAGCGCCAAAGGTGGATGTCTGGAGCAGTGTCT	1981
Db	1973	CTCCGAGGTGGTCTGTGGCAGGAGCTTGCAGCGCCAAAGGTGGATGTCTGGAGCAGTGTCT	2032
Qy	1982	GTATGATGTGCACATGCTCAACGGCTCCACCCCTGACTCAGTCTTCTCCGAGGCCCGC	2041
Db	2033	GTATGATGTGCACATGCTCAACGGCTTGCACCCCTGACTCAGTCTTCTCCGAGGCCCGC	2092
Qy	2042	TCTGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGGCC	2101
Db	2093	TCTGCCTCAAGATTGCCAGGAGCCTCCGCTGTGAGGGAGATCCCAACCTCTCTGGGCC	2152
Qy	2102	CTCTCACAGCCAGGCCATCCAGAGGGGTGAGGAAAGGCCATCCACCGGTGTGTCTG	2161
Db	2153	CTCTCACAGCCAGGCCATCCAGAGGGGTGAGGAAAGGCCATCCACCGGTGTGTCTG	2212
Qy	2162	CAGCGAGCTGGGAGGAGGTTGAACCGGGCACTACAGCAAGTGGGAGTCTCAAGAGCC	2221
Db	2213	CAGCGAGCTGGGAGGAGGTTGAACCGGGCACTACAGCAAGTGGGAGTCTCAAGAGCC	2272
Qy	2222	CTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTCACAC	2281
Db	2273	CTTGGAGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTCACAC	2332
Qy	2282	AGACCTTCCATGCCCGCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCCGCCAGCTG	2341
Db	2333	AGACCTTCCATGCCCGCAGAGAGCTTTTCGCCAAGGGCCCCAGGGCCCCCGCCAGCTG	2392
Qy	2342	AGGAGACAAAGGAGAGACCCCTTAAGCTCAGACCTCTCTTCCACAGAGCCCCCAGAGC	2401
Db	2393	AGGAGACAAAGGAGAGACCCCTTAAGCTCAGACCTCTCTTCCACAGAGCCCCCAGAGC	2452
Qy	2402	CAAAACAAGTCTCTCCCTTGACTTTTGAAGAGGAGTCTGGGATGTGGGAAACCCCTTAC	2461
Db	2453	CAAAACAAGTCTCTCCCTTGACTTTTGAAGAGGAGTCTGGGATGTGGGAAACCCCTTAC	2512
Qy	2462	CTCTGTCTCTCCCTGGAGCCGCCCTGCCAGAAACCCAGCTCACCAGAGCGGAAAGCAA	2521
Db	2513	CTCTGTCTCTCCCTGGAGCCGCCCTGCCAGAAACCCAGCTCACCAGAGCGGAAAGCAA	2572
Qy	2522	CCGTCCGGAGCAGGAACGTGAGCAGCTGGAAATAGAAATTAATCTCTCAAACAGCTGTCCC	2581
Db	2573	CCGTCCGGAGCAGGAACGTGAGCAGCTGGAAATAGAAATTAATCTCTCAAACAGCTGTCCC	2632
Qy	2582	AGCCATTTTCTCTGGAGGACAGAGCAAAATCTCTGTGCTCCCTCAGCATCGACAGCTCT	2641
Db	2633	AGCCATTTTCTCTGGAGGACAGAGCAAAATCTCTGTGCTCCCTCAGCATCGACAGCTCT	2692
Qy	2642	CCCTGTCCGATGACAGTGAAGAAACCCATCAAAGGCTCTCAAAGCTCCGGGACACCC	2701
Db	2693	CCCTGTCCGATGACAGTGAAGAAACCCATCAAAGGCTCTCAAAGCTCCGGGACACCC	2752
Qy	2702	TGAGCTCAGGCGTACATCTCTGGAGCAGCAGGCCGAGGTCTGAAGTCTCAGTGTGAACA	2761
Db	2753	TGAGCTCAGGCGTACATCTCTGGAGCAGCAGGCCGAGGTCTGAAGTCTCAGTGTGAACA	2812

QY	2762	TGTTGTGGCCCGGGGGCGGCCACCGAACATCTTCAATAGTGTTGAAGTTC	2821
DB	2813	TGTTGTCTGGCCGGGGCGGCCACCGACATCTTCAATAGTGTTGAAGTTC	2872
QY	2822	AAATAACAGTCTCTTAATAGTGAACACTGTCACATCCGGAGTTCCACCGGTCAAAGTGG	2881
DB	2873	AAATAACAGTCTCTTAATAGTGAACACTGTCACATCCGGAGTTCCACCGGTCAAAGTGG	2932
QY	2882	GAGACATCGCCACTGGGCATCAGCAGCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACA	2941
DB	2933	GAGACATCGCCACTGGGCATCAGCAGCAGATCCAGCTGCAGCTTTCAGCTTGGTCAACA	2992
QY	2942	AAGACGGCGAGCCTGTTCCGTACGACATGAGAGTGCAGATCTGGGCATCGACTCGAGT	3001
DB	2993	AAGACGGCGAGCCTGTTCCGTACGACATGAGAGTGCAGATCTGGGCATCGACTCGAGT	3052
QY	3002	GCACATCGGCCCTGTATGGCAGCTTTCGCCCTGAGCTTGGAGGTCAAGCATGGCCAGCTGG	3061
DB	3053	GCACATCGGCCCTGTATGGCAGCTTTCGCCCTGAGCTTGGAGGTCAAGCATGGCCAGCTGG	3112
QY	3062	AGAACAGGCCCTAACCCTGCCCCTCACCGCCGGCTCCACATCTGCCGGAAGCAGCTTTC	3121
DB	3113	AGAACAGGCCCTAACCCTGCCCCTCACCGCCGGCTCCACATCTGCCGGAAGCAGCTTTC	3171
QY	3122	TGCTCGGTGCACGATGCTGCCCCGTGAAAAACACAGSCTCAGCCGTTCCACGGGATTCCCA	3181
DB	3172	TGCTCGGTGCACGATGCTGCCCCGTGAAAAACACAGSCTCAGCCGTTCCACGGGATTCCCA	3230
QY	3182	CCCCCGGCTCACAGTGGGAACACAGGCCCTTCGACAGCAGCAAGTGGGGCAAGCAATG	3241
DB	3231	-CCCCCGGCTCAGCAGTGGGAACACAGGCCCTTCGACAGCAGCAAGTGGGGCAAGCAATG	3289
QY	3242	CCTCCACAGGATTTCACACCTGAGCCCTGCCACCTCGTGAAGAAAACHA-TCGCCACCT	3300
DB	3290	CCTCCACAGGATTTCACACCTGAGCCCTGCCACCTCGTGAAGAAAACHA-TCGCCACCT	3349
QY	3301	GAAGAGACAGAAAGGAGATGGCAGGAGTTACCTTGGGGAACAAAAACAGGATCTTTTTCT	3360
DB	3350	GAAGAGACAG-AGGAGGATGGCAGGAGTTACCTTGGGGAACAAAAACAGG---ATCTTCTC	3404
QY	3361	GCCCTCTCTCAGTCAGATTGGCTGACCGCTTGATCAGTGACCATTTGTTGGCAGAC	3420
DB	3405	TGCCCTCTCTCAGTCAGATTGGCTGACCGCTTGATCAGTGACCATTTGTTGGCAGAC	3464
QY	3421	AGGGCAGCAGCTTCCAGCCTGGGTGAGAAGGGGTGGCGAGCCCTTCGCCCTTCACC	3480
DB	3465	AGGGCAGCAGCTTCCAGCCTGGGTGAGAAGGGGTGGCGAGCCCTTCGCCCTTCACC	3524
QY	3481	CTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGCCCAAACTCAGGTTTCAGTCAGAA	3540
DB	3525	CTCCAGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGCCCAAACTCAGGTTTCAGTCAGAA	3584
QY	3541	CCAGGTCAGCAGATTATGCCCGCGTAGTTAAGGGGGCCCTTAAACCCCTTGCTGCG	3600
DB	3585	CCAGGTCAGCAGATTATGCCCGCGTAGTTAAGGGGGCCCTTAAACCCCTTGCTGCG	3644
QY	3601	CTACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTCACCTCGGGAAG	3660
DB	3645	CTACCTGGCCAGCTCACCCCTTTTGGGTGTAGGGGAAAGAAATGCTCACCTCGGGAAG	3704
QY	3661	GCTCCCTGGTAGAATACACACACTTTTCAGGTTGTTGCAACACAGGTTCTGAGTTGACC	3720
DB	3705	GCTCCCTGGTAGAATACACACACTTTTCAGGTTGTTGCAACACAGGTTCTGAGTTGACC	3764
QY	3721	TCTGGTTTCAGCCAAGACCAAGAGGTGTGAAGTGAAGTGGTTCTCAGTCCCCAGACA	3780
DB	3765	TCTGGTTTCAGCCAAGACCAAGAGGTGTGAAGTGAAGTGGTTCTCAGTCCCCAGACA	3824
QY	3781	TGTGCCCTTTGCTGTGGGTACATCTTTCCCAAGCAGCAGAGCCCGAGCCCTTCA	3840
DB	3825	TGTGCCCTTTGCTGTGGGTACATCTTTCCCAAGCAGCAGAGCCCGAGCCCTTCA	3884





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 10, 2005, 09:44:05 ; Search time 1758.13 Seconds  
(without alignments)  
11471.453 Million cell updates/sec

Title: US-09-155-676B-6  
Perfect score: 4596  
Sequence: 1 agcgggggactgtgccgtg.....gatgacaaatgttaaaaaa 4596

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 12985247 seqs, 219411492 residues

Total number of hits satisfying chosen parameters: 25970494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA New:\*

- 1: /cgn2\_6/ptodata/2/pna/pct NEW COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pna/pct NEW COMB.seq2:\*
- 3: /cgn2\_6/ptodata/2/pna/US06 NEW COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/pna/US07 NEW COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pna/US08 NEW COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/pna/US09 NEW COMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq:\*
- 8: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq2:\*
- 9: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq3:\*
- 10: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq4:\*
- 11: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq5:\*
- 12: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pna/US11 NEW COMB.seq3:\*
- 15: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	4596	100.0	4596	6	US-09-155-676B-6
2	4578.4	99.6	4596	8	US-10-450-763-15294
3	4437.6	96.6	4684	15	US-60-659-397-566
4	4243.2	92.3	4458	15	US-60-659-397-567
5	1851.2	40.3	2631	6	US-09-155-676B-3
6	1554.2	33.8	65967	15	US-60-659-397-12129
7	1474.4	32.1	2760	8	US-10-450-763-15295
8	1345.6	29.3	1400	13	US-11-060-756-560
9	1345.6	29.3	1400	13	US-11-060-756-4832
10	676.2	14.7	3658	15	US-60-680-544-32424
11	676.2	14.7	3658	15	US-60-680-473-32424
12	627.6	13.7	722	8	US-10-450-763-15296
13	440.6	9.6	761	15	US-60-680-544-10785
14	440.6	9.6	761	15	US-60-680-544-11056
15	440.6	9.6	761	15	US-60-680-473-10785
16	440.6	9.6	761	15	US-60-680-473-11056
17	200.6	4.4	201	15	US-60-659-397-7882
18	200.6	4.4	201	15	US-60-659-397-7883

19	200.6	4.4	201	15	US-60-659-397-7885	Sequence 7885, Ap
20	200.6	4.4	201	15	US-60-659-397-7886	Sequence 7886, Ap
21	200.6	4.4	201	15	US-60-659-397-7887	Sequence 7887, Ap
22	200.6	4.4	201	15	US-60-659-397-7888	Sequence 7888, Ap
23	200.6	4.4	201	15	US-60-659-397-7889	Sequence 7889, Ap
24	200.6	4.4	201	15	US-60-659-397-7891	Sequence 7891, Ap
25	200.6	4.4	201	15	US-60-659-397-7892	Sequence 7892, Ap
26	200.6	4.4	201	15	US-60-659-397-7893	Sequence 7893, Ap
27	200.6	4.4	201	15	US-60-659-397-32722	Sequence 32722, A
28	200.6	4.4	201	15	US-60-659-397-32724	Sequence 32724, A
29	200.6	4.4	201	15	US-60-659-397-32758	Sequence 32758, A
30	153.2	3.3	201	15	US-60-659-397-7884	Sequence 7884, Ap
31	153.2	3.3	201	15	US-60-659-397-7890	Sequence 7890, Ap
32	153.2	3.3	201	15	US-60-659-397-32757	Sequence 32757, A
33	150	3.3	201	15	US-60-659-397-32751	Sequence 32751, A
34	136.8	3.0	493	1	PCT-US05-00517-813	Sequence 813, App
35	128.6	2.8	201	15	US-60-659-397-32755	Sequence 32755, A
36	119.2	2.6	1353	15	US-60-680-544-37753	Sequence 37753, A
37	119.2	2.6	1353	15	US-60-680-473-37753	Sequence 37753, A
38	98	2.1	201	15	US-60-659-397-32710	Sequence 32710, A
39	88.8	1.9	2207	11	US-10-940-774A-5071	Sequence 5071, Ap
40	88	1.9	201	15	US-60-659-397-32726	Sequence 32726, A
41	83.8	1.8	3089	6	US-09-608-890A-9	Sequence 9, Appli
42	78	1.7	201	15	US-60-659-397-32752	Sequence 32752, A
43	78	1.7	599	9	US-10-972-079-7876	Sequence 7876, Ap
44	75.2	1.6	600	9	US-10-972-079-7875	Sequence 7875, Ap
45	70	1.5	1024	8	US-10-450-763-15297	Sequence 15297, A

ALIGNMENTS

RESULT 1

US-09-155-676B-6  
; Sequence 6, Application US/09155676B  
; GENERAL INFORMATION:  
; APPLICANT: WALLACH, David  
; APPLICANT: MALININ, Nikolai  
; APPLICANT: BOLDIN, Mark  
; APPLICANT: KOVALENKO, Andrei  
; APPLICANT: METT, Igor  
; TITLE OF INVENTION: MODULATORS OF TNF RECEPTOR ASSOCIATED FACTOR (TRAF), THEIR  
; FILE OF INVENTION: PREPARATION AND USE  
; FILE REFERENCE: WALLACH-21  
; CURRENT APPLICATION NUMBER: US/09/155.676B  
; CURRENT FILING DATE: 1999-01-04  
; PRIOR APPLICATION NUMBER: PCT/IL97/00117  
; PRIOR FILING DATE: 1997-04-01  
; PRIOR APPLICATION NUMBER: IL 117800  
; PRIOR FILING DATE: 1996-04-02  
; PRIOR APPLICATION NUMBER: IL 119133  
; PRIOR FILING DATE: 1996-08-26  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 6  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-155-676B-6

Query Match	100.0%;	Score 4596;	DB 6;	Length 4596;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4596;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	AGCGGGGGGACTGTGCGGTGGAACTGTAGCTGTGTAAGGTGGACTCTGTACCATTTG	60	
Db	1	AGCGGGGGGACTGTGCGGTGGAACTGTAGCTGTGTAAGGTGGACTCTGTACCATTTG	60	
Qy	61	AGGATGTTTGAGGATGAGTATGTGGCAGAGGCACATAAACAGGCAGAGACCCCTTT	120	
Db	61	AGGATGTTTGAGGATGAGTATGTGGCAGAGGCACATAAACAGGCAGAGACCCCTTT	120	
Qy	121	GCCCTGCTTTCTCCCAACCAAGGCTGACCTGTGTTCTCCCAAGTCTGGGATCTA	180	

|||||  
Db 121 GCGCTGCTTTCTCCCCAACCCAAAGGCTGACCTGTGTCTCCCAAGTCTGGGATCTA 180  
QY 181 AGTGACCTGCTCTGTGTTGGTCTCTCTCAGATGAGCACAAGCCCTGGGAGATGCGAGTG 240  
Db 181 AGTGACCTGCTCTGTGTTGGTCTCTCTCAGATGAGCACAAGCCCTGGGAGATGCGAGTG 240  
QY 241 ATGGAAATGGCTGCGCAGGTGCCCTGTGCTCAGCAGTGGGCGACGAGAGAACTCCGCC 300  
Db 241 ATGGAAATGGCTGCGCAGGTGCCCTGTGCTCAGCAGTGGGCGACGAGAGAACTCCGCC 300  
QY 301 AAGCCAAAGGAGAGACGCGCCACTTGGGGAAGAAACAGAGCTCGTCTACAAGCTTGAG 360  
Db 301 AAGCCAAAGGAGAGACGCGCCACTTGGGGAAGAAACAGAGCTCGTCTACAAGCTTGAG 360  
QY 361 GCGCTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACC 420  
Db 361 GCGCTGGAGAGAGCCCTGTGTTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACC 420  
QY 421 AAGGCGACAGCCAAAGGAGTCCGAGCGAGGCGCAGCTGCCATCTCTATCATCCGCCAG 480  
Db 421 AAGGCGACAGCCAAAGGAGTCCGAGCGAGGCGCAGCTGCCATCTCTATCATCCGCCAG 480  
QY 481 GCTGAGTGTGAAATAGCAAGTTCAGCCCCACCTTTTTCAGAAACGCAATTTTCATCGCT 540  
Db 481 GCTGAGTGTGAAATAGCAAGTTCAGCCCCACCTTTTTCAGAAACGCAATTTTCATCGCT 540  
QY 541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATTCAGATCCCCCAACAAATGTGGCCCAT 600  
Db 541 GGGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATTCAGATCCCCCAACAAATGTGGCCCAT 600  
QY 601 GCTACAGAGGGCAAAATGCCCGTGTGTGTTGGAAGGAAAGCGTCGACAGAAAGCCCGG 660  
Db 601 GCTACAGAGGGCAAAATGCCCGTGTGTGTTGGAAGGAAAGCGTCGACAGAAAGCCCGG 660  
QY 661 AAGAAACGGAAGAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCTTTGGCCAAA 720  
Db 661 AAGAAACGGAAGAGAGCTCAAAGTCCCTGGCTCATGACGAGTGGCTTTGGCCAAA 720  
QY 721 CCGCTCCCGACAGCCCTGAGCAGAGAGTGCACATCCAGTGGCTCATGACGAGTGGCTTTGGCCAAA 780  
Db 721 CCGCTCCCGACAGCCCTGAGCAGAGAGTGCACATCCAGTGGCTCATGACGAGTGGCTTTGGCCAAA 780  
QY 781 CCACCTCGCGCCCATATGTTAGAAACACCCGCGAGTTTACCAAGCTCTGAAAGAACCA 840  
Db 781 CCACCTCGCGCCCATATGTTAGAAACACCCGCGAGTTTACCAAGCTCTGAAAGAACCA 840  
QY 841 GGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCCGCGCTCTGCTCGA 900  
Db 841 GGCCTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGGCTTACGGCCGCGCTCTGCTCGA 900  
QY 901 TCAGAACTCCAAACTGATGACGCCCTTGGAATGTCTGAAACCAAGTGTGGAACCTGCAAC 960  
Db 901 TCAGAACTCCAAACTGATGACGCCCTTGGAATGTCTGAAACCAAGTGTGGAACCTGCAAC 960  
QY 961 CACCCCGAGGAGGGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTATAGCAGACTG 1020  
Db 961 CACCCCGAGGAGGGCCCTGCGCCCTGCGCCAGCAGCCCTTCCCTATAGCAGACTG 1020  
QY 1021 CCTCATCTCCCTTCCCATTCACCCCTCTCCAGCCCTTGAAACCTTCAACCTCTGAGTCTTC 1080  
Db 1021 CCTCATCTCCCTTCCCATTCACCCCTCTCCAGCCCTTGAAACCTTCAACCTCTGAGTCTTC 1080  
QY 1081 CTGGGCAAACTGGCTGTGTAGACAGCAGCAAGAACCTTGGCTTGACCCACACCTGAGCAAA 1140  
Db 1081 CTGGGCAAACTGGCTGTGTAGACAGCAGCAAGAACCTTGGCTTGACCCACACCTGAGCAAA 1140  
QY 1141 CTGGGCTGTGTAGACAGTCCAAAGCCCTGCGCTGCGCCACACCTGAGCCAGCTGCTG 1200  
Db 1141 CTGGGCTGTGTAGACAGTCCAAAGCCCTGCGCTGCGCCACACCTGAGCCAGCTGCTG 1200  
QY 1201 TCTCGTGTGCCCATGAGAAGTCTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260

Db 1201 TCTCGTGTGCCCATGAGAAGTCTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGC 1260  
QY 1261 AGCGTGAGCTCAAGCCAGGCCCAAGCCTGACCTGAGCCCAAGACCTGGGCGACAGCG 1320  
Db 1261 AGCGTGAGCTCAAGCCAGGCCCAAGCCTGACCTGAGCCCAAGACCTGGGCGACAGCG 1320  
QY 1321 GGCTCCAGATCCCGGGAGGCCCAAGCCTGAGGAACAACGAGGGTGTCTGCTCACT 1380  
Db 1321 GGCTCCAGATCCCGGGAGGCCCAAGCCTGAGGAACAACGAGGGTGTCTGCTCACT 1380  
QY 1381 GAGAACTCAAGCCAGTGTATGATPACGAGAAAGTCCACTGGGCGACAGCACAG 1440  
Db 1381 GAGAACTCAAGCCAGTGTATGATPACGAGAAAGTCCACTGGGCGACAGCACAG 1440  
QY 1441 CTCCGCTGGGCGAGGCTCTTCCGAGAGTGCACAGATGGAGGAACAAGACTGGC 1500  
Db 1441 CTCCGCTGGGCGAGGCTCTTCCGAGAGTGCACAGATGGAGGAACAAGACTGGC 1500  
QY 1501 TTCCAGTCCGCTGTCAAAAAGGCTCGCTGGAAGTATTTCCGGCGACAGAGCTGATGGCA 1560  
Db 1501 TTCCAGTCCGCTGTCAAAAAGGCTCGCTGGAAGTATTTCCGGCGACAGAGCTGATGGCA 1560  
QY 1561 TGTGAGGATTCACCTCAAGAAATTCCTTTGTATGAGTGTGAGAAAGGGCT 1620  
Db 1561 TGTGAGGATTCACCTCAAGAAATTCCTTTGTATGAGTGTGAGAAAGGGCT 1620  
QY 1621 TGGGTCAACATCTTCATGAGAGTGTGAAAGTGGCTCCCTGGGCGACAGTGGTCAAGGAG 1680  
Db 1621 TGGGTCAACATCTTCATGAGAGTGTGAAAGTGGCTCCCTGGGCGACAGTGGTCAAGGAG 1680  
QY 1681 CAGGCTGTCTCCCAAGAGGACCGGCGCTGTACTACTCTGGGCGAGGCTGGAGGCTG 1740  
Db 1681 CAGGCTGTCTCCCAAGAGGACCGGCGCTGTACTACTCTGGGCGAGGCTGGAGGCTG 1740  
QY 1741 GAATACCTCCACTCAAGAAAGTTCGATCGGGAAGTCAAGCTGCAAAAGTGTCTCTG 1800  
Db 1741 GAATACCTCCACTCAAGAAAGTTCGATCGGGAAGTTCGATCGGGAAGTCAAAAGTGTCTCTG 1800  
QY 1801 TCCAGCGATGGAGGCCACGACGCTCTGTGACTTTGGCCATGCTGTGCTTCAACCT 1860  
Db 1801 TCCAGCGATGGAGGCCACGACGCTCTGTGACTTTGGCCATGCTGTGCTTCAACCT 1860  
QY 1861 GATGCTTGGGAAAGTCTTCTCACAGGAGTACATCCCTGGGCGACAGAGCCACATG 1920  
Db 1861 GATGCTTGGGAAAGTCTTCTCTCACAGGAGTACATCCCTGGGCGACAGAGCCACATG 1920  
QY 1921 GCTCGGAGTGGTCTGGGCGAGAGTGCAGACGCAAGGTGATGTCTGAGAGAGTGC 1980  
Db 1921 GCTCGGAGTGGTCTGGGCGAGAGTGCAGACGCAAGGTGATGTCTGAGAGAGTGC 1980  
QY 1981 TGTATGATGTGCATGCTCAAGCGGCTGCGACCCCTGGACTCAGTTCTTCGAGGGCG 2040  
Db 1981 TGTATGATGTGCATGCTCAAGCGGCTGCGACCCCTGGACTCAGTTCTTCGAGGGCG 2040  
QY 2041 CTCTGCTTCAAGATTGGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGCC 2100  
Db 2041 CTCTGCTTCAAGATTGGCCAGGAGCTCCGCTGTGAGGAGATCCCACTCTCTGCGCC 2100  
QY 2101 CCTCTCAGAGCCAGGCTCAAGAGGGGCTGAGAAAGAGCCCATCAAGCGCTGTCT 2160  
Db 2101 CCTCTCAGAGCCAGGCTCAAGAGGGGCTGAGAAAGAGCCCATCAAGCGCTGTCT 2160  
QY 2161 GCAGGGAGCTGGGAGGAAAGTGAACCGGCGACTACAGAGTGGGAGGCTGAGAGC 2220  
Db 2161 GCAGGGAGCTGGGAGGAAAGTGAACCGGCGACTACAGAGTGGGAGGCTGAGAGC 2220  
QY 2221 CTTTGGAGGGAGAAATATAAAGAACCAAGACATCCACCGCCCAAAATCAAGCCAAATACCAC 2280  
Db 2221 CTTTGGAGGGAGAAATATAAAGAACCAAGACATCCACCGCCCAAAATCAAGCCAAATACCAC 2280  
QY 2281 CAGACCTTCCATGCGCCAGCGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCGAGCT 2340  
Db 2281 CAGACCTTCCATGCGCCAGCGAGAGCTTTTCGCAAGGGCCCGAGGGCCCGGCGAGCT 2340

Qy	2341	GAGGAGACAA	CAGG	CAGAGCCCTTAAG	CTCAG	CCTCCTCTCTCC	CAAGAGCC	CCCCCAGAG	2400
Db	2341	GAGGAGACAA	CAGG	CAGAGCCCTTAAG	CTCAG	CCTCCTCTCTCC	CAAGAGCC	CCCCCAGAG	2400
Qy	2401	CCAAACAGTCT	CTCCTT	GACATTTGAG	CAAGGAGTCTGG	GATGCTGG	AAACCCCTTA	2460	
Db	2401	CCAAACAGTCT	CTCCTT	GACATTTGAG	CAAGGAGTCTGG	GATGCTGG	AAACCCCTTA	2460	
Qy	2461	CCTCTGCTCCT	CCCTG	AGCCAGCCCTG	CCAGAAACCC	AGCTCACC	AGAGCGGAAAGCA	2520	
Db	2461	CCTCTGCTCCT	CCCTG	AGCCAGCCCTG	CCAGAAACCC	AGCTCACC	AGAGCGGAAAGCA	2520	
Qy	2521	ACCGTCCGGAG	CAAGGA	CTCAG	AGCTGGAATATTT	CTCAAC	CAGAGCCTGTCC	2580	
Db	2521	ACCGTCCGGAG	CAAGGA	CTCAG	AGCTGGAATATTT	CTCAAC	CAGAGCCTGTCC	2580	
Qy	2581	CAGCCATTTT	CTCTG	AGGAGCAGG	CAAAATTTCT	CTGCTCCT	CAGCATCGAGCCTC	2640	
Db	2581	CAGCCATTTT	CTCTG	AGGAGCAGG	CAAAATTTCT	CTGCTCCT	CAGCATCGAGCCTC	2640	
Qy	2641	TCCCTGTCGAT	GACAGT	GAGAAAC	CCATCAAGG	CCTCTCAAG	CTCGCGGACACC	2700	
Db	2641	TCCCTGTCGAT	GACAGT	GAGAAAC	CCATCAAGG	CCTCTCAAG	CTCGCGGACACC	2700	
Qy	2701	CTGAGCTCAG	CGGTAC	ACTCTCTG	AGCAGCGCGG	CGCTCGAAG	CTCAGCTGGAAC	2760	
Db	2701	CTGAGCTCAG	CGGTAC	ACTCTCTG	AGCAGCGCGG	CGCTCGAAG	CTCAGCTGGAAC	2760	
Qy	2761	ATGTGTGCGC	CGCGCGG	CGCCAC	CGACACCC	CAAGCTCT	CAATGTGTGAAGTC	2820	
Db	2761	ATGTGTGCGC	CGCGCGG	CGCCAC	CGACACCC	CAAGCTCT	CAATGTGTGAAGTC	2820	
Qy	2821	CAAAATACAGT	CTCTTAA	TGTTGA	ACCTGACATCGG	AGTTCC	ACCGGCTCAAAAGTG	2880	
Db	2821	CAAAATACAGT	CTCTTAA	TGTTGA	ACCTGACATCGG	AGTTCC	ACCGGCTCAAAAGTG	2880	
Qy	2881	GGAGACATCG	CACTG	GCATCAG	CAGCCAGAT	CCCAGCT	TCAGCTTGGTCAAC	2940	
Db	2881	GGAGACATCG	CACTG	GCATCAG	CAGCCAGAT	CCCAGCT	TCAGCTTGGTCAAC	2940	
Qy	2941	AAAGCGGCGA	GCCTG	TTCGATAG	CATGAGGTG	CCAGAT	CTCGGCGATCGACCTGGAG	3000	
Db	2941	AAAGCGGCGA	GCCTG	TTCGATAG	CATGAGGTG	CCAGAT	CTCGGCGATCGACCTGGAG	3000	
Qy	3001	TGCACATGCG	CCCTG	ATGAGT	CTGAGCTG	GGAGTCA	AGCATGGCCAGCTG	3060	
Db	3001	TGCACATGCG	CCCTG	ATGAGT	CTGAGCTG	GGAGTCA	AGCATGGCCAGCTG	3060	
Qy	3061	GAGAACAGG	CCCTA	ACCTG	CCCTCC	ACCGCGG	CTCCACATGCGGAAAGCAGCCTTC	3120	
Db	3061	GAGAACAGG	CCCTA	ACCTG	CCCTCC	ACCGCGG	CTCCACATGCGGAAAGCAGCCTTC	3120	
Qy	3121	CTGCTCGGTG	CAAGAT	GTCCCTG	AAACACAGG	CTCAG	CCGTTCCAGGGGATTGCCA	3180	
Db	3121	CTGCTCGGTG	CAAGAT	GTCCCTG	AAACACAGG	CTCAG	CCGTTCCAGGGGATTGCCA	3180	
Qy	3181	CCCCCCCCG	CTCAGT	GGGAAC	CAAGGCT	CGCAG	CAAGGCTGGGGGCAAGCAAT	3240	
Db	3181	CCCCCCCCG	CTCAGT	GGGAAC	CAAGGCT	CGCAG	CAAGGCTGGGGGCAAGCAAT	3240	
Qy	3241	GCCTTCCAG	GATTTT	CACACCT	GAGCCCTG	CCCACT	GTGAAACATCCGCCACGT	3300	
Db	3241	GCCTTCCAG	GATTTT	CACACCT	GAGCCCTG	CCCACT	GTGAAACATCCGCCACGT	3300	
Qy	3301	GAAGAGACAG	AAGGAGT	TGACAG	AGTTACTCTG	GGGAAAC	CAAAACAGGGATCTTTTCT	3360	
Db	3301	GAAGAGACAG	AAGGAGT	TGACAG	AGTTACTCTG	GGGAAAC	CAAAACAGGGATCTTTTCT	3360	
Qy	3361	CCCCCTGCT	CCAGT	CGAGT	TGGCCTG	ACCCCTG	GATCAGTGACCATTTGTTGGCAGAC	3420	
Db	3361	CCCCCTGCT	CCAGT	CGAGT	TGGCCTG	ACCCCTG	GATCAGTGACCATTTGTTGGCAGAC	3420	

Qy	3421	AGGGGAGAG	CAGCTT	CCAGCCT	GGGT	CAGAA	GGGTGGG	CGAGCCCTTCG	CCCCCTCACC	3480
Db	3421	AGGGGAGAG	CAGCTT	CCAGCCT	GGGT	CAGAA	GGGTGGG	CGAGCCCTTCG	CCCCCTCACC	3480
Qy	3481	CTCCAGGCTC	GTGTG	AGAGTGT	CAAGTGT	TAA	GGGCCCC	CAAACTC	CAGGTTCA	3540
Db	3481	CTCCAGGCTC	GTGTG	AGAGTGT	CAAGTGT	TAA	GGGCCCC	CAAACTC	CAGGTTCA	3540
Qy	3541	CCAGGT	CAG	CAGGTAT	ATCC	CCCGCGT	AGTTAA	AGGGGGCCCT	CTTAA	3600
Db	3541	CCAGGT	CAG	CAGGTAT	ATCC	CCCGCGT	AGTTAA	AGGGGGCCCT	CTTAA	3600
Qy	3601	CTCACCTGG	CAGCT	CACCCCT	TTTGG	GTAG	GGGAAAG	ATGCCT	CGACCCCTGGG	3660
Db	3601	CTCACCTGG	CAGCT	CACCCCT	TTTGG	GTAG	GGGAAAG	ATGCCT	CGACCCCTGGG	3660
Qy	3661	GCTCCCTGG	TAG	AAATAC	CAC	CACACT	TTTCA	GGTGTG	TGCAACAC	3720
Db	3661	GCTCCCTGG	TAG	AAATAC	CAC	CACACT	TTTCA	GGTGTG	TGCAACAC	3720
Qy	3721	TCTGGTT	CAG	CAAGGA	CCAA	AGAGTGT	TAA	GTGAAGTGT	TCTCAGT	3780
Db	3721	TCTGGTT	CAG	CAAGGA	CCAA	AGAGTGT	TAA	GTGAAGTGT	TCTCAGT	3780
Qy	3781	TGTGCCCT	TTTGT	CTGCT	GGCTA	CCACT	CTTCCC	CAGAGCAG	CGCCCGC	3840
Db	3781	TGTGCCCT	TTTGT	CTGCT	GGCTA	CCACT	CTTCCC	CAGAGCAG	CGCCCGC	3840
Qy	3841	GGCCAG	CACTG	CCCCC	AGACT	CGCT	GGC	ACTCAGT	TCCCTCAT	3900
Db	3841	GGCCAG	CACTG	CCCCC	AGACT	CGCT	GGC	ACTCAGT	TCCCTCAT	3900
Qy	3901	GATG	CAG	GATAT	GCTT	CAC	AGGAAC	AGTCT	GTGAT	3960
Db	3901	GATG	CAG	GATAT	GCTT	CAC	AGGAAC	AGTCT	GTGAT	3960
Qy	3961	CAG	CAG	AGAGAG	AGT	CCG	GGCCCC	CAGCCCA	CTATCAGT	4020
Db	3961	CAG	CAG	AGAGAG	AGT	CCG	GGCCCC	CAGCCCA	CTATCAGT	4020
Qy	4021	CAG	AG	CAC	AGT	CAC	CTG	CACTC	ACCTG	4080
Db	4021	CAG	AG	CAC	AGT	CAC	CTG	CACTC	ACCTG	4080
Qy	4081	TGCC	CG	AG	CG	CACTT	GTG	CACTT	CAAG	4140
Db	4081	TGCC	CG	AG	CG	CACTT	GTG	CACTT	CAAG	4140
Qy	4141	GGC	AG	GGC	AG	GGC	AG	GGC	AG	4200
Db	4141	GGC	AG	GGC	AG	GGC	AG	GGC	AG	4200
Qy	4201	AC	AGT	CTT	GAG	CTG	TC	CAAT	GTG	4260
Db	4201	AC	AGT	CTT	GAG	CTG	TC	CAAT	GTG	4260
Qy	4261	AT	AG	CA	CCCTT	CC	CACTT	GG	CTT	4320
Db	4261	AT	AG	CA	CCCTT	CC	CACTT	GG	CTT	4320
Qy	4321	CT	CA	CGT	AG	CA	TTAA	AT	TC	4380
Db	4321	CT	CA	CGT	AG	CA	TTAA	AT	TC	4380
Qy	4381	GG	GC	AG	GG	GA	CG	CG	AG	4440
Db	4381	GG	GC	AG	GG	GA	CG	CG	AG	4440
Qy	4441	CT	TT	GT	CC	AG	CT	GT	CA	4500
Db	4441	CT	TT	GT	CC	AG	CT	GT	CA	4500
Qy	4501	CAC	ATA	AG	AA	GA	AACT	TC	TT	4560
Db	4501	CAC	ATA	AG	AA	GA	AACT	TC	TT	4560

Db 4501 CACATAAAGAACTGCAGCCCTTGGTACTGCAGAGTCTGGGTCTAGAGAACTCTTTGTAAG 4560  
QY 4561 CAATAAAGTTGGGTGATGACAAATGTTAAAAAA 4596  
Db 4561 CAATAAAGTTGGGTGATGACAAATGTTAAAAAA 4596  
RESULT 2  
US-10-450-763-15294  
; Sequence 15294, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; PRIOR FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 15294  
; LENGTH: 4596  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1117)..(1581)  
; OTHER INFORMATION: 100% homologous to Homo sapiens sperm protein 10, accession  
; OTHER INFORMATION: number M82968, Smith-Waterman Score=809.  
US-10-450-763-15294  
Query Match 99.6%; Score 4578.4; DB 8; Length 4596;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4590; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
QY 5 GGGGACTGTGCGTGTGAAAGTGTAGCTGTGTA-AGTGAATCTGTACCATTGAGG 63  
Db 5 GGGGACTGTGCGTGTGAAAGTGTAGCTGTGAGGTGACTCTGTACCATTGAGG 64  
QY 64 ATGTTGGAGTAGTAGTGTGTCAGAGGCACACATAAAGCAGCAGACCTTTGCC 123  
Db 65 ATGTTGGAGTAGTAGTGTGTCAGAGGCACACATAAAGCAGCAGACCTTTGCC 124  
QY 124 CTGCTCTTTCTCCCCCAACCCCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT 183  
Db 125 CTGCTCTTTCTCCCCCAACCCCAAGCTGACCTGTGTTCTCCAGGTCTGGGATTTCTAAGT 184  
QY 184 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCAAGCTGGGAGATGGCAGTGATG 243  
Db 185 GACCTGCTCTGTGTTGTTGTTCTCTCAGGATGAGCAAGCTGGGAGATGGCAGTGATG 244  
QY 244 GAAATGGCTGCCAGGTGCCCTTGCTCAGCAGTGGGCGACAGAGAACTCCCCAAG 303  
Db 245 GAAATGGCTGCCAGGTGCCCTTGCTCAGCAGTGGGCGACAGAGAACTCCCCAAG 304  
QY 304 CCAAGAGGAGAACGCCGCCACTGGGGAGAAACAGAGCTCCGTTCTACAAGCTTGAGGCC 363  
Db 305 CCAAGAGGAGAACGCCGCCACTGGGGAGAAACAGAGCTCCGTTCTACAAGCTTGAGGCC 364  
QY 364 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGAAGTATACCAAG 423  
Db 365 GTGGAGAGAGCCCTGTGTTCTGCGGAAAGTGGGAGATCCTGAATGAAGTATACCAAG 424  
QY 424 GGCAGAGCAAGAGAGCTCCAGGCGAGGCTGCCATCTCTATCATCCCGAGGCT 483  
Db 425 GGCAGAGCAAGAGAGCTCCAGGCGAGGCTGCCATCTCTATCATCCCGAGGCT 484  
QY 484 GAGTGTGAGAAATAGCCAGAGTTACGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 543

Db 485 GAGTGTGAGAAATAGCCAGAGTTCAGCCCCACCTTTTCAGAACGCAATTTTCATCGCTGGG 544  
QY 544 TCCAAAAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCT 603  
Db 545 TCCAAAAGTACAGCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCT 604  
QY 604 ACAGAGGCGAAATGGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 663  
Db 605 ACAGAGGCGAAATGGCCCGTGTGTTGGAGGAAAGCGTCGAGCAAAAGCCCGGAAG 664  
QY 664 AAAACGAAAGAAAGAGCTCAAAGTCCCTGCTCATGCAAGAGTGGCTTTGGCCAAACCC 723  
Db 665 AAAACGAAAGAAAGAGCTCAAAGTCCCTGCTCATGCAAGAGTGGCTTTGGCCAAACCC 724  
QY 724 CTCCCCAGGACCCCTGAGCAGAGAGTGCACCAATCCCAAGTGCAGGAGATGAGTCTCA 783  
Db 725 CTCCCCAGGACCCCTGAGCAGAGAGTGCACCAATCCCAAGTGCAGGAGATGAGTCTCA 784  
QY 784 CTGCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCCTCTGAAGAAACAGGC 843  
Db 785 CTGCGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACAGCCCTCTGAAGAAACAGGC 844  
QY 844 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTTACGCGCCGCTCTGCTCGATCA 903  
Db 845 CTTGGGCAACTCTGTTTAAAGCAGCTTGGCGAGGCGCTTACGCGCCGCTCTGCTCGATCA 904  
QY 904 GAACTCCACAACTGATCAGCCCTTGGCAATGTCTGAACCAAGTGTGAAACTGCACAC 963  
Db 905 GAACTCCACAACTGATCAGCCCTTGGCAATGTCTGAACCAAGTGTGAAACTGCACAC 964  
QY 964 CCCAGGACGAGGCGCCCTGCGCCCGCAGCAGCCCTTCCCTATAGCAGACTGCT 1023  
Db 965 CCCAGGACGAGGCGCCCTGCGCCCGCAGCAGCCCTTCCCTATAGCAGACTGCT 1024  
QY 1024 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCTCTGGAGTCTTCTCTG 1083  
Db 1025 CATCCCTTCCCATTTCCACCTCTCCAGCCCTGGAACCTCACCTCTGGAGTCTTCTCTG 1084  
QY 1084 GGCMAACTGGCTGTGTAGACAGCAGAAACCTTTGCTGACCCACACCTCAGCAAACTG 1143  
Db 1085 GGCMAACTGGCTGTGTAGACAGCAGAAACCTTTGCTGACCCACACCTCAGCAAACTG 1144  
QY 1144 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGCT 1203  
Db 1145 GCCTGTGTAGACAGTCCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCTGCT 1204  
QY 1204 CGTGTGCTCCATGAGAAAGTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGCAGC 1263  
Db 1205 CGTGTGCTCCATGAGAAAGTCTGTGGAGAAATACCTAGTGCATGCTCTGCAAGGCAGC 1264  
QY 1264 GTGAGCTCAAGCCAGGCCCAAGCTGACCTGAGCCAGAGCTGGGCGAGCAGGGGC 1323  
Db 1265 GTGAGCTCAAGCCAGGCCCAAGCTGACCTGAGCCAGAGCTGGGCGAGCAGGGGC 1324  
QY 1324 TCCAGATCCCGGAGCCAGCCCAAAACTCAGAGCAACAGGGGTGCTCTGCTCACTGAG 1383  
Db 1325 TCCAGATCCCGGAGCCAGCCCAAAACTCAGAGCAACAGGGGTGCTCTGCTCACTGAG 1384  
QY 1384 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCATGCGGCCACGCAACAGCTC 1443  
Db 1385 AAATCAAGCCAGTGGATTATGAGTACCGAGAAAGTCCATGCGGCCACGCAACAGCTC 1444  
QY 1444 GCGCTGGGCGAGAGCTCTTCGGAGAGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1503  
Db 1445 GCGCTGGGCGAGAGCTCTTCGGAGAGTGCACAGGATGGAGGACAAAGCAGACTGGCTTC 1504  
QY 1504 CAGTGCCTGTCAAAGAGGTGCGCTGGAAGTATTTTCGGGCGAGAGAGCTGATGCACTGT 1563  
Db 1505 CAGTGCCTGTCAAAGAGGTGCGCTGGAAGTATTTTCGGGCGAGAGAGCTGATGCACTGT 1564  
QY 1564 GCAGGATGACCTCAACCCAGAAATGTCCTTTGATGAGTGTGAGAGAGGGCTTGG 1623

Db 1565 GCAGGATTGACCTCACCCAGAAATTGTCTCCCTTTGTATGAGAGCTGTGAGAGAGGCGCTTGG 1624  
Qy 1624 GTCAACATCTTTATGAGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1683  
Db 1625 GTCAACATCTTTATGAGAGCTCTGGAAGGTGGCTCCCTGGGCCAGCTGGTCAAGGAGCAG 1684  
Qy 1684 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTTGGGCCAGGCCCTGGAGGGTCTGGAA 1743  
Db 1685 GGCTGTCTCCAGAGGACCGGGCCCTGTACTACCTTGGGCCAGGCCCTGGAGGGTCTGGAA 1744  
Qy 1744 TACCTCCACTCACCAAGATTCTCATGGGAGCTCAAGCTGACAAGCTGCAACGTGCTCCTGTCC 1803  
Db 1745 TACCTCCACTCACCAAGATTCTCATGGGAGCTCAAGCTGACAAGCTGCAACGTGCTCCTGTCC 1804  
Qy 1804 AGCATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT 1863  
Db 1805 AGCATGGGAGCCACGAGCCCTCTGTGACTTTGGCCATGTGTGTCTTTCAACCTGAT 1864  
Qy 1864 GGCTGGGAAAGTCTTGTCTCAGAGGGACTACATCTCCTTGGCA CAGAGACCACATGGCT 1923  
Db 1865 GGCTGGGAAAGTCTTGTCTCAGAGGGACTACATCTCCTTGGCA CAGAGACCACATGGCT 1924  
Qy 1924 CCGAGGTGGTGTGGGAGGAGCTGCGACGCCAAGGTGGATGTCTGGAGCAGCTGTGT 1983  
Db 1925 CCGAGGTGGTGTGGGAGGAGCTGCGACGCCAAGGTGGATGTCTGGAGCAGCTGTGT 1984  
Qy 1984 ATGATGTGACATGCTCAACGGCTGCGACGCCCTGGACTCAGTTCTTCCGAGGGCCGCTC 2043  
Db 1985 ATGATGTGACATGCTCAACGGCTGCGACGCCCTGGACTCAGTTCTTCCGAGGGCCGCTC 2044  
Qy 2044 TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTCTGCGCCCT 2103  
Db 2045 TGCCTCAAGATTGCCAGGAGCTCCGCTGTGAGGAGATCCACCCCTCTCTGCGCCCT 2104  
Qy 2104 CTACAGCCAGGCCATTCAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2163  
Db 2105 CTACAGCCAGGCCATTCAGAGGGGCTGAGGAAAGAGCCCATCCACCGGTGTCTGCA 2164  
Qy 2164 GCGAGCTGGAGGGAAGGTGAACCGGSCACTACAGCAAGTGGGAGTCTGAAGAGCCCT 2223  
Db 2165 GCGAGCTGGAGGGAAGGTGAACCGGSCACTACAGCAAGTGGGAGTCTGAAGAGCCCT 2224  
Qy 2224 TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG 2283  
Db 2225 TGGAGGGGAGAAATATAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATACCAACAG 2284  
Qy 2284 ACCCTCATGCCAGCCAGAGAGCTTTGGCAAGGGCCCGAGGGCCCGGCCAGCTGAG 2343  
Db 2285 ACCCTCATGCCAGCCAGAGAGCTTTGGCAAGGGCCCGAGGGCCCGGCCAGCTGAG 2344  
Qy 2344 GAGACAAGCAGAGCCCTTAAGCTCCAGCTCCTCTCCACAGAGCCCGCCAGAGCCCA 2403  
Db 2345 GAGACAAGCAGAGCCCTTAAGCTCCAGCTCCTCTCCACAGAGCCCGCCAGAGCCCA 2404  
Qy 2404 AACAAAGTCTCCTCCCTTGTACTTTGAGCAAGAGGAGTCTGGGATGTGGGAAACCTTACCT 2463  
Db 2405 AACAAAGTCTCCTCCCTTGTACTTTGAGCAAGAGGAGTCTGGGATGTGGGAAACCTTACCT 2464  
Qy 2464 CTGTCTCCTCGAGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGAACCC 2523  
Db 2465 CTGTCTCCTCGAGAGCCAGCCCTTGCAGAAACCCAGCTCACAGAGCGGAAAGAACCC 2524  
Qy 2524 GTCCGGAGCAGGAATCTGACAGCTGGAAATAGAAATTTCTCTCAACAGCTGTCCAG 2583  
Db 2525 GTCCGGAGCAGGAATCTGACAGCTGGAAATAGAAATTTCTCTCAACAGCTGTCCAG 2584  
Qy 2584 CCATTTTCTCGAGGAGCAGGACAAATCTCTGTGCTCTCAGCATCGAGCTCTCC 2643  
Db 2585 CCATTTTCTCGAGGAGCAGGACAAATCTCTGTGCTCTCAGCATCGAGCTCTCC 2644  
Qy 2644 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCCCGGACACCCCTG 2703  
Db 2645 CTGTCCGATGACAGTGAAGAAACCCATCAAGGCCCTCTCAAGCTCCCGGACACCCCTG 2704

Qy 2704 AGCTCAGGCGTACACTCTCTGGAGCAGCCAGCGAGCTCGAAGCTCCAGCTCGAACAATG 2763  
Db 2705 AGCTCAGGCGTACACTCTCTGGAGCAGCCAGCGAGCTCGAAGCTCCAGCTCGAACAATG 2764  
Qy 2764 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGTAAGTCCAA 2823  
Db 2765 GTGCTGGCCCGGGGGCGGCCACCGACACCCCAAGCTATTTCATGTGTGTAAGTCCAA 2824  
Qy 2824 ATACAGTCTCTTAATGTGAAACCTGACATCCGGAGTTCACACCGGTCAAAAGTGGGA 2883  
Db 2825 ATACAGTCTCTTAATGTGAAACCTGACATCCGGAGTTCACACCGGTCAAAAGTGGGA 2884  
Qy 2884 GACATGCCCACTGGCATCAGCAGCCAGATCCAGCTCGAGCTTTCAGCTTGGTCAACAAA 2943  
Db 2885 GACATGCCCACTGGCATCAGCAGCCAGATCCAGCTCGAGCTTTCAGCTTGGTCAACAAA 2944  
Qy 2944 GACGGGAGCCTGTTGCTACGATGGAGTGGCAGACTCGGGCATCGAGCTCGAGTGC 3003  
Db 2945 GACGGGAGCCTGTTGCTACGATGGAGTGGCAGACTCGGGCATCGAGCTCGAGTGC 3004  
Qy 3004 ACATGGCCCTCATGAGCAGCTTCGCTGAGCTGGAGGTCAAGCATGGCCAGCTGGAG 3063  
Db 3005 ACATGGCCCTCATGAGCAGCTTCGCTGAGCTGGAGGTCAAGCATGGCCAGCTGGAG 3064  
Qy 3064 AACAGGCCCTAACCTGSCCTCCACCGCCGCTCCACACTGCCGGAAGCAGCTTCTCTG 3123  
Db 3065 AACAGGCCCTAACCTGSCCTCCACCGCCGCTCCACACTGCCGGAAGCAGCTTCTCTG 3124  
Qy 3124 CTGCTGCAAGTGTGCTCCCTGAAACACAGGCTCAGCGTTCCAGGGGATTCAGAGCC 3183  
Db 3125 CTGCTGCAAGTGTGCTCCCTGAAACACAGGCTCAGCGTTCCAGGGGATTCAGAGCC 3184  
Qy 3184 CCGCGCTCAGTGGGAAACAGGGCTCCAGCAGCAAGTGGGGCAGAGCAAGTGC 3243  
Db 3185 CCGCGCTCAGTGGGAAACAGGGCTCCAGCAGCAAGTGGGGCAGAGCAAGTGC 3244  
Qy 3244 TCCAGAGATTTCACACTGAGCCCTTGCCCACTGCTGTAAGAAACATCCGCCACGTGAA 3303  
Db 3245 TCCAGAGATTTCACACTGAGCCCTTGCCCACTGCTGTAAGAAACATCCGCCACGTGAA 3304  
Qy 3304 GAGACAGAGGAGTGGCAGGATTTACCTGGGGAACAAACAGGGATCTTTTCTGTC 3363  
Db 3305 GAGACAGAGGAGTGGCAGGATTTACCTGGGGAACAAACAGGGATCTTTTCTGTC 3364  
Qy 3364 CTTGCTCCAGTGGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGAGCAGG 3423  
Db 3365 CTTGCTCCAGTGGCTGACCCGCTTGGATCAGTGACCATTTGTTGGCAGAGCAGG 3424  
Qy 3424 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGGCCCTCACCCCTC 3483  
Db 3425 GGAGAGCAGCTTCCAGCCTGGGTGAGAGGGGTGGGCGAGCCCTTCGGGCCCTCACCCCTC 3484  
Qy 3484 CAGGCTCTGTGAGAGTCAAGTGTGTAAGGGCCCAAACTCAGGTTTCAAGTGAAGACCA 3543  
Db 3485 CAGGCTCTGTGAGAGTCAAGTGTGTAAGGGCCCAAACTCAGGTTTCAAGTGAAGACCA 3544  
Qy 3544 GGTTCAGAGTATGCCCGCGGTAGGTTAAGGGGCCCTCTTAACCCCTTCGCTGGCCTC 3603  
Db 3545 GGTTCAGAGTATGCCCGCGGTAGGTTAAGGGGCCCTCTTAACCCCTTCGCTGGCCTC 3604  
Qy 3604 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCTGACCTTGGGAAAGCT 3663  
Db 3605 ACCTGGCCAGCTCACCCCTTTGGGTGAGGGGAAAGAAATGCTGACCTTGGGAAAGCT 3664  
Qy 3664 CCCTGGTAGAATACACACTTTTTCAGGTTGTTGCAACACAGCTCTGAGTTGACCTCT 3723  
Db 3665 CCCTGGTAGAATACACACTTTTTCAGGTTGTTGCAACACAGCTCTGAGTTGACCTCT 3724  
Qy 3724 GGTTCAGCAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCTCCCGACACATGT 3783  
Db 3725 GGTTCAGCAGGACCAAGAGGTGTGTAAGTGAAGTGGTTCTCAGTCTCCCGACACATGT 3784

```
QY 3784 GCCCTTTTGTCTGCTGCTACCACTCTTCCCCAGAGCAGCAGCCCGCCAGCCCTTCAGGC 3843
Db 3785 GCCCTTTTGTCTGCTGCTACCACTCTTCCCCAGAGCAGCAGCCCGCCAGCCCTTCAGGC 3844
QY 3844 CCAGCAGTCGCCCGCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTGAT 3903
Db 3845 CCAGCAGTCGCCCGCAGACTCGCTGGCACTCAGTTCCTCATCTGTAAAGGTGAAGGTGAT 3904
QY 3904 CGAGGATATGCTGTGACAGAAACAGTCTGTGGATGACATGATCAGTCTTAAGGAAGCAG 3963
Db 3905 CGAGGATATGCTGTGACAGAAACAGTCTGTGGATGACATGATCAGTCTTAAGGAAGCAG 3964
QY 3964 CAGAGAGAGACGTCGCCGCGCCCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCCCCAG 4023
Db 3965 CAGAGAGAGACGTCGCCGCGCCCGCCAGCCCACTATCAGTGTCCAGCGTGTGTTCCCCAG 4024
QY 4024 AGCAGAGTCAGCATCACACTGACATCACTCACTGCTGCCCTGCCCTGCCAGAGGGTACTGC 4083
Db 4025 AGCAGAGTCAGCATCACACTGACACTCACTGCTGCCCTGCCCTGCCAGAGGGTACTGC 4084
QY 4084 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTGCCAGGGC 4143
Db 4085 CGACGGCACTTTGCACTCTGATGACCTCAAGCACTTTTCATGGCTGCCCTGCCAGGGC 4144
QY 4144 AGGGCAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACA 4203
Db 4145 AGGGCAGGGCAGTGCACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGAAGGACACA 4204
QY 4204 GTCTTGAGCTGCCACATGATGATGATCTCCTCAAACTCTTCCAGATTTCTTAAGAATA 4263
Db 4205 GTCTTGAGCTGCCACATGATGATGATCTCCTCAAACTCTTCCAGATTTCTTAAGAATA 4264
QY 4264 GCACCCCTTCCCACTGCCCCAGCTTAGCTCTTCTCCAGGGAGCTACTCAGACTC 4323
Db 4265 GCACCCCTTCCCACTGCCCCAGCTTAGCTCTTCTCCAGGGAGCTACTCAGACTC 4324
QY 4324 ACCTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCTCAACCTCCTGGG 4383
Db 4325 ACCTAGCATTAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCTCAACCTCCTGGG 4384
QY 4384 CGAGGGGACCGCAGACTCCGTGGGAGAGCTCATTCOCACATCTTGGCAAGACGCTT 4443
Db 4385 CGAGGGGACCGCAGACTCCGTGGGAGAGCTCATTCOCACATCTTGGCAAGACGCTT 4444
QY 4444 TGTCAGCTGTCCATTTAGTGCAGTCTCTCCGGGGAGAGCCCGGCCCCAGCAC 4503
Db 4445 TGTCAGCTGTCCATTTAGTGCAGTCTCTCCGGGGAGAGCCCGGCCCCAGCAC 4504
QY 4504 ATAAAGAACTGCAGCCTTGTTACTGAGAGTCTGGTGTGAGAGAACTCTTTGTAAGCAA 4563
Db 4505 ATAAAGAACTGCAGCCTTGTTACTGAGAGTCTGGTGTGAGAGAACTCTTTGTAAGCAA 4564
QY 4564 TAAAGTTGGGTGATGACAAATGTTAAAAA 4595
Db 4565 TAAAGTTGGGTGATGACAAATGTTAAAAA 4596
```

```
RESULT 3
US-60-659-397-566
; Sequence 566, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 566
```

```
; LENGTH: 4684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-659-397-566

Query Match          96.6%; Score 4437.6; DB 15; Length 4684;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 4543; Conservative 6; Mismatches 30; Indels 11; Gaps 8;

QY 3 CGGGGGGAGCTGTGCCGTGTGGAAACGTGTAGCTGTGTA-AGGTGGAAGCTGTGTACCAATTGA 61
Db 53 CGGGGGGAGCTGTGCCGTGTGGAAACGTGTAGCTGTGTGAGAGGTGGAGCTCTGTGTACCAATTGA 112
QY 62 GGAATGTTGGAGGATGAGTATGTGTGGCAGAGGCACATAAAACAGGCGAGAGACCTTTTG 121
Db 113 GGAATGTTGGAGGATGAGTATGTGTGGCAGAGGCACATAAAACAGGCGAGAGACCTTTTG 172
QY 122 CCCCTGCCCTTTCTCCCCCAACCCCAAGGCTGACCTGTGTCTCTCCCAAGGCTCTGGGATTTCTAA 181
Db 173 CCCCTGCCCTTTCTCCCCCAACCCCAAGGCTGACCTGTGTCTCTCCCAAGGCTCTGGGATTTCTAA 232
QY 182 GTGACCTCTCTGTGTTTGTGTTCTCTCAGGATGAGCAAAAGCTTGGGAGATGGCAGTGA 241
Db 233 GTGACCTCTCTGTGTTTGTGTTCTCTCAGGATGAGCAAAAGCTTGGGAGATGGCAGTGA 292
QY 242 TGGAAATGGCCTGCCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAGGAACCTCCCA 301
Db 293 TGGAAATGGCCTGCCCAGGTGCCCTGGCTCAGCAGTGGGGCAGCAGAGGAACCTCCCA 352
QY 302 AGCCAAAGGAGAGACGCGCCACCTGGGGAAAGAAAACAGAGCTCCGTCTACAAGCTTTGAGG 361
Db 353 AAGCAAGGAGAGAGAGCGCGCCACCTGGGGAAAGAAAACAGAGCTCCGTCTACAAGCTTTGAGG 412
QY 362 CCGTGGAAAGAGAGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACCA 421
Db 413 CCGTGGAAAGAGAGCCCTGTGTCTCGGAAAGTGGGAGATCCTGAATGACGTGATTACCA 472
QY 422 AGGCGACAGCCAAAGGAAGGCTCCGAGGCGAGGCGCAGCTGCCATCTCTATCATCGCCAGG 481
Db 473 AGGCGACAGCCAAAGGAAGGCTCCGAGGCGAGGCGCAGCTGCCATCTCTATCATCGCCAGG 532
QY 482 CTGAGTGTGAGATAGCAAGAGTTTACGCCCAACCTTTTCAGAAACGCAATTTTCATCGCTG 541
Db 533 CTGAGTGTGAGATAGCAAGAGTTTACGCCCAACCTTTTCAGAAACGCAATTTTCATCGCTG 592
QY 542 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATG 601
Db 593 GGTCCAAACAGTACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAAATGTGGCCCATG 652
QY 602 CTACAGAGGGCAAAATGSCCGTGTGTGGAAGGGAAAGCGTCGACAGCAAAAGCCCGGA 661
Db 653 CTACAGAGGGCAAAATGSCCGTGTGTGGAAGGGAAAGCGTCGACAGCAAAAGCCCGGA 712
QY 662 AGAAACGGAAGAAGAGAGCTCAAAAGTCCCTGGCTCATGACAGAGTGGCCTTTGGCCAAAC 721
Db 713 AGAAACGGAAGAAGAGAGCTCAAAAGTCCCTGGCTCATGACAGAGTGGCCTTTGGCCAAAC 772
QY 722 CCCTCCCGCAGGACCCCTGAGCAGGAGAGCTGCAACATCCAGTGCAGAGGATGAGTCTC 781
Db 773 CCCTCCCGCAGGACCCCTGAGCAGGAGAGCTGCAACATCCAGTGCAGAGGATGAGTCTC 832
QY 782 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACCAAGCCTCTGAAGGAACCAAG 841
Db 833 CACTCGGCGCCCATATGTTAGAAACACCCCGCAGTTTCAACCAAGCCTCTGAAGGAACCAAG 892
QY 842 GCCTTGGGCAACTCTGTGTTTAAAGCAGAGCTTGGCAGGGGCTTACGGCGGCTCTGCCTCGAT 901
Db 893 GCCTTGGGCAACTCTGTGTTTAAAGCAGAGCTTGGCAGGGGCTTACGGCGGCTCTGCCTCGAT 952
QY 902 CAGAACTCCAAACTGATCAGCCCTTGGATGTCTGAAACCACTGTGGAAACTGGACC 961
Db 953 CAGAACTCCAAACTGATCAGCCCTTGGCAATGTCTGAACCACTGTGGAAACTGGACC 1012
```



QY 962 ACCCCAGGAGCGGCCCTGCCCCCTGCCAGCACCCCTTCCCTATAGCAGCTGC 1021  
Db 1013 ACCCCAGGAGCGAGGCCCTGCCCCCTGCCAGCACCCCTTCCCTATAGCAGCTGC 1072  
QY 1022 CTATCCCTTCCCATTTCCACCTCTTCCAGCCCTGGAACCTCACCCCTCTGAGTCTCTCC 1081  
Db 1073 CTATCCCTTCCCATTTCCACCTCTTCCAGCCCTGGAACCTCACCCCTCTGAGTCTCTCC 1132  
QY 1082 TGGGCAAACTGGCTGTGTAGACGCCAGAAACCCCTTGCCTTGACCCACCTGAGCAAAAC 1141  
Db 1133 TGGGCAAACTGGCTGTGTAGACGCCAGAAACCCCTTGCCTTGACCCACCTGAGCAAAAC 1192  
QY 1142 TGGCTGTGTAGACAGTCAAAAGCCCTGCTGGCCACACCTGAGGCCAGCTGCTGT 1201  
Db 1193 TGGCTGTGTAGACAGTCAAAAGCCCTGCTGGCCACACCTGAGGCCAGCTGCTGT 1252  
QY 1202 CTGTGTGTGCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1261  
Db 1253 CTGTGTGTGCCATGAGAAATTTCTGTGGAGGAATACCTAGTGCATGCTCTGCAAGGCA 1312  
QY 1262 GCGTAGCTCAAGCCAGGCCCCACAGCCTTGACAGCCTGGCCCAAGACCTTGGGCAGCAGGG 1321  
Db 1313 GCGTAGCTCAAGCCAGGCCCCACAGCCTTGACAGCCTGGCCCAAGACCTTGGGCAGCAGGG 1372  
QY 1322 GCTCCAGATCCGGGAGCCAGCCCAAACTGAGGACCAAGAGGCTGCTGCTCACTG 1381  
Db 1373 GCTCCAGATCCGGGAGCCAGCCCAAACTGAGGACCAAGAGGCTGCTGCTCACTG 1432  
QY 1382 AGAAACTCAAGCCAGTGTATGATACCGAGAGAAAGTCCACTTGGGCCAGCAGCAGC 1441  
Db 1433 AGAAACTCAAGCCAGTGTATGATACCGAGAGAAAGTCCACTTGGGCCAGCAGCAGC 1492  
QY 1442 TCCGCTTGGGAGAGGCTCTTCCGAGAGGTGCAAGGATGGAGGACCAAGCAGACTGGCT 1501  
Db 1493 TCCGCTTGGGAGAGGCTCTTCCGAGAGGTGCAAGGATGGAGGACCAAGCAGACTGGCT 1552  
QY 1502 TCCAGTGGCTCAAAAAGTGGCCCTGGAAGTATTTCCGGGCAGAGAGCTGATGCAT 1561  
Db 1553 TCCAGTGGCTCAAAAAGTGGCCCTGGAAGTATTTCCGGGCAGAGAGCTGATGCAT 1612  
QY 1562 GTGCAGGATTCACCTCACCCAGAAATGTCCTTTGTATGGAGCTGTGAGAGAGGGCTT 1621  
Db 1613 GTGCAGGATTCACCTCACCCAGAAATGTCCTTTGTATGGAGCTGTGAGAGAGGGCTT 1672  
QY 1622 GGGTCAACATCTTATGGAGCTGTGGAAGGTGGCTCCCTGGGCAGCTGTCTCAAGAGC 1681  
Db 1673 GGGTCAACATCTTATGGAGCTGTGGAAGGTGGCTCCCTGGGCAGCTGTCTCAAGAGC 1732  
QY 1682 AGGGTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGCCAGGCCCTGAGAGGTCTGG 1741  
Db 1733 AGGGTGTCTCCAGAGGACCGGGCCCTGTACTACTGGGCCAGGCCCTGAGAGGTCTGG 1792  
QY 1742 AATACCTCCACTCAGAGGATTCGATGGGAGCTCAAGCTGACAACTGCTCCTGT 1801  
Db 1793 AATACCTCCACTCAGAGGATTCGATGGGAGCTCAAGCTGACAACTGCTCCTGT 1852  
QY 1802 CCAGCGATGGAGGACCAAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCTG 1861  
Db 1853 CCAGCGATGGAGGACCAAGCAGCCCTCTGTGACTTTGGCCATGCTGTGCTTCAACCTG 1912  
QY 1862 ATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCCAGAGACCCACATGG 1921  
Db 1913 ATGGCTGGGAAAGTCTTGTCTCAGGGGACTACATCCCTGGCCAGAGACCCACATGG 1972  
QY 1922 CTCCGAGGTGGTCTGGGAGGAGCTGCGAGCCCAAGGTGGATGTCTGGAGCAGCTGCT 1981  
Db 1973 CTCCGAGGTGGTGTCTGGGAGGAGCTGCGAGCCCAAGGTGGATGTCTGGAGCAGCTGCT 2032  
QY 1982 GTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTCTCCGAGGGCCGC 2041  
Db 2033 GTATGATGCTGCATGCTCAACGGCTGCCACCCCTGGACTCAGTCTCTCCGAGGGCCGC 2092  
QY 2042 TCTGCCCTCAAGATTTGCCAGGAGCCTCCGGCTGTGAGGGAGATCCCACTCTCTGCGCCC 2101

Db 2093 TCTGCCCTCAAGATTTGCCAGGAGCCTCCGGCTGTGAGGGAGATCCCACTCTCTGCGCCC 2152  
QY 2102 CTCTCACAGGCCAGGCCATCCAGAGGGGTGAGGAAAGAGCCCATCCACCGCGTGTCTG 2161  
Db 2153 CTCTCACAGGCCAGGCCATCCAGAGGGGTGAGGAAAGAGCCCATCCACCGCGTGTCTG 2212  
QY 2162 CAGCGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCC 2221  
Db 2213 CAGCGAGCTGGGAGGGAAGGTGAACCGGGCACTACAGCAAGTGGGAGGTCTGAAGAGCC 2272  
QY 2222 CTTGAGGGGAGATATAAAGAACCAAGACATCCACCGCCAAATCAAGCCAAATTACCACC 2281  
Db 2273 CTTGAGGGGAGATATAAAGAACCAAGAVATCCACCGCCAAATCAAGCCAAATTACCACC 2332  
QY 2282 AGACCTCTCCATGCCAGCCAGAGAGCTTTCCGCCAAGGGGCCAGGGCCCCGGCCAGCTG 2341  
Db 2333 AGACCTCTCCATGCCAGCCAGAGAGCTTTCCGCCAAGGGGCCAGGGCCCCGGCCAGCTG 2392  
QY 2342 AGGAGACAAAGGAGAGGCCCTTAAGCTCCAGCCTCTCTCCACCAAGAGCCCCCAGAGC 2401  
Db 2393 AGGAGACAAAGGAGAGGCCCTTAAGCTCCAGCCTCTCTCCACCAAGAGCCCCCAGAGC 2452  
QY 2402 CAACAAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTAC 2461  
Db 2453 CAACAAAGTCTCTCTCCCTTGAATTTGAGCAAGGAGGAGTCTGGGATGTGGGAAACCTTTAC 2512  
QY 2462 CTCTGTCTCTCTGGAGCCAGGCCCTCCAGAAACCCAGCTCACAGAGCGGAAAGCAA 2521  
Db 2513 CTCTGTCTCTCTGGAGCCAGGCCCTCCAGAAACCCAGCTCACAGAGCGGAAAGCAA 2572  
QY 2522 CCGTCCCGAGCAGGAACTGCAGCAGCTGGAATATTAATCTCAACAGCCTGTCCC 2581  
Db 2573 CCGTCCCGAGCAGGAACTGCAGCAGCTGGAATATTAATCTCAACAGCCTGTCCC 2632  
QY 2582 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCTGCTGCTCAGCATCGACAGCCTCT 2641  
Db 2633 AGCCATTTTCTCTGGAGGAGCAGGAGCAAAATCTCTCTGCTGCTCAGCATCGACAGCCTCT 2692  
QY 2642 CCCTGTCCGATGACAGTGAGAAAGAACCCATCAAGGGCTCTCAAGAGCTCGGGGACACCC 2701  
Db 2693 CCCTGTCCGATGACAGTGAGAAAGAACCCATCAAGGGCTCTCAAGAGCTCGGGGACACCC 2752  
QY 2702 TGAGCTCAGCGCTACACTCTCTGGAGCAGCAGGCCGAGGCTCGAAGCTCAGCTTGAAGACA 2761  
Db 2753 TGAGCTCAGCGCTACACTCTCTGGAGCAGCAGGCCGAGGCTCGAAGCTCAGCTTGAAGACA 2812  
QY 2762 TGGTGTGGCCCGGGGGCGGCCCAACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCC 2821  
Db 2813 TGGTGTGGCCCGGGGGCGGCCCAACCGACACCCCAAGCTATTTCAATGGTGTGAAAGTCC 2872  
QY 2822 AATATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGTCAAAGTGG 2881  
Db 2873 AATATACAGTCTCTTAATGGTGAACACCTGCACATCCGGGAGTTCCACCGGTCAAAGTGG 2932  
QY 2882 GAGACATTCGCCACTGGCATCAGCAGCAGATCCCAGCTGAGCCTTCAGCTTGTCTACCA 2941  
Db 2933 GAGACATTCGCCACTGGCATCAGCAGCAGATCCCAGCTGAGCCTTCAGCTTGTCTACCA 2992  
QY 2942 AAGACGGGAGCCTGTTTGGCTACGATGAGGTGTCAGACTCGGGCATCGACCTGCACT 3001  
Db 2993 AAGACGGGAGCCTGTTTGGCTACGATGAGGTGTCAGACTCGGGCATCGACCTGCACT 3052  
QY 3002 GCACACTGGCCCTGATGGCAGCTTCCCTTGGAGCTGAGGGGTCAAGCATGGCCAGCTGG 3061  
Db 3053 GCACACTGGCCCTGATGGCAGCTTCCCTTGGAGCTGAGGGGTCAAGCATGGCCAGCTGG 3112  
QY 3062 AGAACAGGCCCTAACCTGCTGCCCTCCACCGCGGCTCCACACTTCCCGGAAAGCAGCTTCC 3121  
Db 3113 AGAACAGGCCCTAACCTGCTGCCCTCCACCGCGGCTCCACACTTCCCGGAAAGCAGCTTCC 3171  
QY 3122 TGCTCGGTGACGATGCTGCCCTTGAAACACAGGCTCAGCGCTTCCAGGGGATTCGAC 3181

3172	TGCTCGGTGCA	CGATGCTGCCCTTG-AAAACA	CAGGCTCAGCCGTTTC	CAGGGATCTGCCAG	3233
3182	CCCCCGGCTCA	CAGTGGGAAC	CAGGSCCTCG	CAGCAGCAAGTGGGGG	CAAGCAATG
3231	-CCCCGGCTC	AGTGGGAC	CAGGSCCTCG	CAGCAGCAAGTGGGGG	CAAGCAATG
3242	CTCTCCAGGAT	TTTCA	CACCTGAGCC	TGCCCAACCTTGCTG	AAAAAACA-TCGCGCA
3290	CTCTCCAGGAT	TTTCA	CACCTGAGCC	TGCCCAACCTTGCTG	TGAAAAACA
3301	GAAGAGACAA	AGGAGATG	CGCAGGAGTT	ACCTGGGGAAC	AAAAACAGGGATCTTTTCT
3350	GAAGAGACAG	-AGGAGATG	CGCAGGAGTT	ACCTGGGGAAC	AAAAACAGG-ATCTTCTC
3361	GC	CCCTGCTCC	AGTTCGAGTT	GGCCTCAGCCGCTT	TGGATCAGTACCA
3405	TG	CCCTGCTCC	AGTTCGAGTT	GGCCTGACCCGCTT	TGGATCAGTACCA
3421	AG	GGGAGAG	CAGCTTC	CAGCCTGGGT	CAGAAGGGGTGGCGAGCC
3465	AG	GGGAGAG	CAGCTTC	CAGCCTGGGT	CAGAAGGGGTGGCGAGCC
3481	CT	CCAGGCTG	CTG	GAGTGTCA	AGTGTGTAAGGGCCCAA
3525	CT	CAGGCTG	CTG	GAGTGTCA	AGTGTGTAAGGGCCCAA
3541	CC	AGGTCAG	CAGGATAT	CGCCGCGCT	AGGTTAAAGGGGCGCCCT
3585	CA	AGGTCAG	CAGGATAT	CGCCGCGCT	AGGTTAAAGGGGCGCCCT
3601	CT	CACCTGG	CCAGCTCA	CCCTTTTGGGT	GTAGGGGAAAAAGATG
3645	CT	CACCTGG	CCAGCTCA	CCCTTTTGGGT	GTAGGGGAAAAAGATG
3661	GC	TCCCTGG	TAGAAATAC	CACACACTTT	TTCAGTTGTGAA
3705	GCT	CCCTGGT	TAGAAATAC	CACACACTTT	TTCAGTTGTGAA
3721	TC	TGGTTAC	CCAGGACAA	AGAGGTGT	AGTGAAGTGTCTC
3765	TC	TGGTTAC	CCAGGACAA	AGAGGTGT	AGTGAAGTGTCTC
3781	TG	TGCCCTTT	TGCTGGCT	TACCACTCTT	CCCCAGACAGCGCCCGAGCCCTTCA
3825	TG	TGCCCTTT	TGCTGGCT	TACCACTCTT	CCCCAGACAGCGCCCGAGCCCTTCA
3841	GG	CCAGCACT	CGCC	CAGACTCG	TGGCACTCAGTTCCCTCATCTG
3885	GG	CCAGCACT	CGCC	CAGACTCG	TGGCACTCAGTTCCCTCATCTG
3901	GAT	GCAGGATAT	GCCTC	ACAGGAA	CAGTCTGTGATGGA
3945	GAT	GCAGGATAT	GCCTC	ACAGGAA	CAGTCTGTGATGGA
3961	CAG	CAGAGAGAG	AGC-TCC	GGGCGCC	CCAGCCCACTCATAGTTC
4005	CAG	CAGAGAGAG	AGC-TCC	GGGCGCC	CCAGCCCACTCATAGTTC
4020	CC	AGAGACA	AGCTCAG	CACTGACA	CTCACCTTGCCTCGCC
4065	CC	AGAGACA	AGCTCAG	CACTGACA	CTCACCTTGCCTCGCC
4080	CT	CCGACG	CACTTTTG	CACTCTGAT	CACTCAAGCACTTT
4125	CT	CCGACG	CACTTTTG	CACTCTGAT	CACTCAAGCACTTT
4140	GG	GAGGGC	CAGGGCAGT	GCATGT	AGGAGCATAGCAAG
4185	GG	GAGGGC	CAGGGCAGT	GCATGT	AGGAGCATAGCAAG
4200	CAC	AGTCTT	CAGCTGT	CCACAT	GCATGTACTCTTCA
4245	CAC	AGTCTT	CAGCTGT	CCACAT	GCATGTACTCTTCA

## RESULT 4

US-60-659-397-567

US-60-633-397-387  
: Sequence 567. Application IIS/60659397

; sequence 567, Application: GENERAL INFORMATION.

;  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL. Michele

;; APPLICANT: CARGILL, Michele  
: APPLICANT: CHANG sheng-Yung

APPLICANT: CHANG, Sheng-Yung  
TITLE OF INVENTION. GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
: TITLE OF INVENTION: RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C

; TITLE OF INVENTION: , RESPONSE TO INTERFERON TREATMENT IN HEPATITIS C  
FIELD OF INVENTION: VIRUS INFECTED SUBJECTS METHODS OF DETECTION AND USES

; TITLE OF INVENTION; VIRUS-I

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001470

; CURRENT APPLICATION NUMBER: US/6

; CURRENT FILING DATE: 2005-03-09

; NUMBER OF SEQ

; SOFTWARE: Fast

; SEQ ID NO 56

; LENGTH: 4458

; TYPE: DNA

ORGANISM: Homo sapiens

US-60-659-397-567

Query Match 92.3%; Score 4243.2; DB 15; Length 4458;

Best Local Similarity 98.7%; Pred. No. 0;

Matches	Conservative	Mismatches	Indels	Gaps
4343	6	39	10	7

RECEIVED  
COMMUNICATIONS SECTION  
FEBRUARY 1967

OV 194 GTGTTGGTCTCTCTCAGGATGAGCACAGCCCTGGGAGATGGCAGTGATGGAAATGGCCT 253

QY

Db 69 GAGATCGGGCGCTTGGCCGGATGAGCACACAGCCTGGGAGATGGCAGTGATGGAATGGCCT 128

DD 69 GAGATCGGCGCTTGGCCGGATGAGCACAAGCCTGGAGAGATGGCAGTGTATGGATTGGCCCTTCTT

254 G C C C A G G T G C C C C T G G C T A G C A G T G G G G C A G C A G G A A C T C C C C A A G C C A A A G G A G A 313

QY 254 GCCAGGIGCCCTGGCTCAGCAGTGGGCAAGCAGHAGGHAICTCCCAHGCCAHHAGGAGA 313

128 CCCCAGCTCCCCCTCGCTCAGCAGTGGGCAGCAGAGGAAGTCCCCAAAGCCAGGAGA 188

Db 369 ATAGCCAAAGAGTTCAGCCCCACCTTTTTCAGAAACGCAATTTTCATCGCTGGGTCCAAACAGT 428  
Qy 554 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCTACAGAGGCA 613  
Db 429 ACAGCCAGTCCGAGAGTCTTGATCAGATCCCAACAATGTGGCCCATGCTACAGAGGCA 488  
Qy 614 AAATGGCCCGTGTGTGGAGGGAAGCCTCGCAGCAAAAGCCCGGAGAAACGGAAGA 673  
Db 489 AAATGGCCCGTGTGTGGAGGGAAGCCTCGCAGCAAAAGCCCGGAGAAACGGAAGA 548  
Qy 674 AGAAGAGCTCAAAAGTCCCTGGCTCATGCAGAGTGGCTTGGCCAAACCCCTCCCAAGGA 733  
Db 549 AGAAGAGCTCAAAAGTCCCTGGCTCATGCAGAGTGGCTTGGCCAAACCCCTCCCAAGGA 608  
Qy 734 CCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGGATGATCTCCACTCGCGCGCC 793  
Db 609 CCCCTGAGCAGGAGAGTGCACCATCCAGTGCAGGAGGATGATCTCCACTCGCGCGCC 668  
Qy 794 CATATGTTAGAAACACCCCGCAGTTCAACCAAGCTCTGMAAGGAACCAAGGCTTGGGCAC 853  
Db 669 CATATGTTAGAAACACCCCGCAGTTCAACCAAGCTCTGMAAGGAACCAAGGCTTGGGCAC 728  
Qy 854 TCTGTTTAAAGCAGTTGGCGAGGCGCTTACGGCCGGCTCTGCCCTCGATCAGAACTCCACA 913  
Db 729 TCTGTTTAAAGCAGTTGGCGAGGCGCTTACGGCCGGCTCTGCCCTCGATCAGAACTCCACA 788  
Qy 914 AACTGATCAGCCCTTGCATGTCTGAACACAGTGTGGAACCTGCACACCCCGCAGGACG 973  
Db 789 AACTGATCAGCCCTTGCATGTCTGAACACAGTGTGGAACCTGCACACCCCGCAGGACG 848  
Qy 974 GAGGCCCTGCGCTGCGCAGGACACCGCTTCCCTATAGCAGACTGGCTCATCCCTTCC 1033  
Db 849 GAGGCCCTGCGCTGCGCAGGACACCGCTTCCCTATAGCAGACTGGCTCATCCCTTCC 908  
Qy 1034 CATTCACCTCTCCAGCCCTGGAAACCTCAACCTCTGGAGTCTCTCTGGGCAAACTGG 1093  
Db 909 CATTCACCTCTCCAGCCCTGGAAACCTCAACCTCTGGAGTCTCTCTGGGCAAACTGG 968  
Qy 1094 CTTGTGTAGACAGCAGAAACCTTTGGCTGACCCACACCTGAGCAAACTGGCTGTGTAG 1153  
Db 969 CTTGTGTAGACAGCAGAAACCTTTGGCTGACCCACACCTGAGCAAACTGGCTGTGTAG 1028  
Qy 1154 ACAGTCAAAGCCCTGCTGGCCCAACACTGTGAGCCAGCTGCTGTCTGTGGTGCC 1213  
Db 1029 ACAGTCAAAGCCCTGCTGGCCCAACACTGTGAGCCAGCTGCTGTCTGTGGTGCC 1088  
Qy 1214 ATGAGAGTCTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGAGCGTGTAGCTCA 1273  
Db 1089 ATGAGAGTCTTCTGTGAGGAATACCTAGTGCATGCTCTGCAAGGAGCGTGTAGCTCA 1148  
Qy 1274 GCCAGGCCCAAGCCTGACAGCCTGGCCAAAGACTTGGCAGACACGGGGCTCCAGATCCC 1333  
Db 1149 GCCAGGCCCAAGCCTGACAGCCTGGCCAAAGACTTGGCAGACACGGGGCTCCAGATCCC 1208  
Qy 1334 GGGAGCCAGCCCAAACTGAGACAAACGAGGTGTCTGCTCACTGAGAAACTCAAGC 1393  
Db 1209 GGGAGCCAGCCCAAACTGAGACAAACGAGGTGTCTGCTCACTGAGAAACTCAAGC 1268  
Qy 1394 CAGTGGATTATGATCGAGAAAGTCCACTGGGCCACGACACAGCTCGGCTGGGCA 1453  
Db 1269 CAGTGGATTATGATCGAGAAAGTCCACTGGGCCACGACACAGCTCGGCTGGGCA 1328  
Qy 1454 GAGGCTCTCTCGAGAGGTGCACAGGATGGAGGAACAGCAGCTGGCTTCCAGTGGCTG 1513  
Db 1329 GAGGCTCTCTCGAGAGGTGCACAGGATGGAGGAACAGCAGCTGGCTTCCAGTGGCTG 1388  
Qy 1514 TCAAAAAGGTGCGCTGGAAGTATTTCCGGCAGAGGAGCTGATGGCATGTGCAGATTGA 1573  
Db 1389 TCAAAAAGGTGCGCTGGAAGTATTTCCGGCAGAGGAGCTGATGGCATGTGCAGATTGA 1448  
Qy 1574 CTTCAACAGATTTGCTCTTGTATGAGCTGTGAGAGAGGCGCTTGGGTCAACATCT 1633  
Db 1449 CTTCAACAGATTTGCTCTTGTATGAGCTGTGAGAGAGGCGCTTGGGTCAACATCT 1508

Qy 1634 TCATGAGCTGCTGGAAGGTGGTCTCCTGGGCCAGCTGGTCAAGGAGCAGGGCTGTCTCC 1693  
Db 1509 TCATGAGCTGCTGGAAGGTGGTCTCCTGGGCCAGCTGGTCAAGGAGCAGGGCTGTCTCC 1568  
Qy 1694 CAGAGGACCGGCCCTGCTACTACTGGGCCAGGCCCTGGAGGGTCTGGAACTACCTCACT 1753  
Db 1569 CAGAGGACCGGCCCTGCTACTACTGGGCCAGGCCCTGGAGGGTCTGGAACTACCTCACT 1628  
Qy 1754 CACGAAGGATTTGCTATGCGGAGCGTCAAAAGCTGCAACAGCTGCTCTGTCACAGCGATGGGA 1813  
Db 1629 CACGAAGGATTTGCTATGCGGAGCGTCAAAAGCTGCAACAGCTGCTCTGTCACAGCGATGGGA 1688  
Qy 1814 GCCACGAGCCCTCTGTGAATTTGGCCATGCTGTGTCTTCAACTGATGGCTTGGGAA 1873  
Db 1689 GCCACGAGCCCTCTGTGTGACTTTGGCCATGCTGTGTCTTCAACTGATGGCTTGGGAA 1748  
Qy 1874 AGTCTTGTCTACAGGGGAGCTACATCCCTGGCAGAGAGCCCATGGCTCCGGAGGTGG 1933  
Db 1749 AGTCTTGTCTACAGGGGAGCTACATCCCTGGCAGAGAGCCCATGGCTCCGGAGGTGG 1808  
Qy 1934 TGCTGGCAGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGCTGC 1993  
Db 1809 TGCTGGCAGAGCTGCGACGCGCAAGGTGGATGTCTGGAGCAGCTGCTGTATGATGCTGC 1868  
Qy 1994 ACATGCTCAAAGGTGGCTGAGGAGAGTCAAGTCTTCCGAGGGCCGCTTGCCTCAAGA 2053  
Db 1869 ACATGCTCAAAGGTGGCTGAGGAGAGTCAAGTCTTCCGAGGGCCGCTTGCCTCAAGA 1928  
Qy 2054 TTGCCAGGAGCTTCCGCTGTGAGGAGATCCACCTCTCTGCGCCCTCTCAAGGCC 2113  
Db 1929 TTGCCAGGAGCTTCCGCTGTGAGGAGATCCACCTCTCTGCGCCCTCTCAAGGCC 1988  
Qy 2114 AGGCCATCAAAGGGGCTGAGGAAAGAGCCATCCACCGCTGTCTGACGCGAGCTGG 2173  
Db 1989 AGGCCATCAAAGGGGCTGAGGAAAGAGCCATCCACCGCTGTCTGACGCGAGCTGG 2048  
Qy 2174 GAGGAAAGGTGAACCGGGCTACTAGCAAGTGGGAGTCTGAAGAGCCCTTTGGAGGGAG 2233  
Db 2049 GAGGAAAGGTGAACCGGGCTACTAGCAAGTGGGAGTCTGAAGAGCCCTTTGGAGGGAG 2108  
Qy 2234 AATATAAGAACCCAGACATCCACCGCAATTAAGGCAATTAACAGAGCCCTCCATG 2293  
Db 2109 AATATAAGAACCCAGAYATCCACCGCAATTAAGGCAATTAACAGAGCCCTCCATG 2168  
Qy 2294 CCAGCGCAGAGAGCTTTCCGCAAGGGCCCGCAGGGCCCGGCGCAGCTGAGGAGCAACAG 2353  
Db 2169 CCAGCGCAGAGAGCTTTCCGCAAGGGCCCGCAGGGCCCGGCGCAGCTGAGGAGCAACAG 2228  
Qy 2354 GCAGAGCCCTTAAGCTCCAGCTCTCTCCACCCAGAGCCCCCAGAGCCCAACAAAGTCTC 2413  
Db 2229 GCAGAGCCCTTAAGCTCCAGCTCTCTCCACCCAGAGCCCCCAGAGCCCAACAAAGTCTC 2288  
Qy 2414 CTCCCTTGAATTTGAGCAAGGAGAGTCTGGATGTGGGAACCCCTTACTCTGTCTCTCC 2473  
Db 2289 CTCCCTTGAATTTGAGCAAGGAGAGTCTGGATGTGGGAACCCCTTACTCTGTCTCTCC 2348  
Qy 2474 TGGAGCCAGCCCTGCGCAGAAACCCAGAGCTCACAGAGCGGAAGCAACCGTCCCGGAGC 2533  
Db 2349 TGGAGCCAGCCCTGCGCAGAAACCCAGAGCTCACAGAGCGGAAGCAACCGTCCCGGAGC 2408  
Qy 2534 AGGAACCTGCAGCAGCTGGAATATAGAAATTAATTCCTCAACAGCTCTCCAGCCATTTTCTC 2593  
Db 2409 AGGAACCTGCAGCAGCTGGAATATAGAAATTAATTCCTCAACAGCTCTCCAGCCATTTTCTC 2468  
Qy 2594 TGGAGGAGCAGGAGCAAAATTTCTGTGCTCTGAGATCGACAGCTCTCCCTGTGCGATG 2653  
Db 2469 TGGAGGAGCAGGAGCAAAATTTCTGTGCTCTGAGATCGACAGCTCTCCCTGTGCGATG 2528  
Qy 2654 ACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCTGAGCTCAGGCG 2713  
Db 2529 ACAGTGAAGAAACCCATCAAGGCTCTCAAGCTCGCGGACACCTGAGCTCAGGCG 2588



;; PRIOR FILING DATE: 1996-04-02  
;; PRIOR APPLICATION NUMBER: IL 119133  
;; PRIOR FILING DATE: 1996-08-26  
;; NUMBER OF SEQ ID NOS: 22  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 3  
;; LENGTH: 2631  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1081)..(1081)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1102)..(1102)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1120)..(1120)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1125)..(1125)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1129)..(1129)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1135)..(1135)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1146)..(1146)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1170)..(1170)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1180)..(1180)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1188)..(1188)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1208)..(1208)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1239)..(1239)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1248)..(1249)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1278)..(1278)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1297)..(1297)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1310)..(1310)  
;; OTHER INFORMATION: n is a, c, g, or t

;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1322)..(1322)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1345)..(1345)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1409)..(1409)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1423)..(1423)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1445)..(1445)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1452)..(1452)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1459)..(1459)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1478)..(1478)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1498)..(1498)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1507)..(1508)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1520)..(1520)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1534)..(1534)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1540)..(1540)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1546)..(1546)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1557)..(1557)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1713)..(1713)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1895)..(1895)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1900)..(1900)  
;; OTHER INFORMATION: n is a, c, g, or t  
;; FEATURE:



```
Db 1501 GGTAAANNNGGGGCGCTCTNAAACCCCTTGCTNGGCGCTNACCTNGGCGCAGCTCANCCT 1560
Qy 3621 CTTTGTGGGTGAGGGGAAAAGAAATGCTGTAGACCTCGGGAAGGCT-CCCTGGTGTAGATACAC 3679
Db 1561 CTTTGTGGGTGAGGGGAAAAGAAATGCTGTAGACCTCGGGAAGGCTWCCCTGGTGTAGATACAC 1620
Qy 3680 CACACTTTTTCAGGTTGTTGCAACACAGAGTCTGTAGTTGACCTCTGTTTTCAGCCAGAGACC 3739
Db 1621 CACACTTTTTCAGGTTGTTGCAACACAGAGTCTGTAGTTGACCTCTGTTTTCAGCCAGAGACC 1680
Qy 3740 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTCTGCTG 3798
Db 1681 AAAGAAGGTGTGAAGTGAAGTGTCTCAGT-CCCCAGACATGTGCCCTTTCTGCTG 1740
Qy 3799 GCTACCACTCTTCCCCAGAGCAGAGCCCGCCGAGCCCTTTTTCAGGCCAGCAGCTGCCCCAG 3858
Db 1741 GCTACCACTCTTCCCCAGAGCAGAGCCCGCCGAGCCCTTTTTCAGGCCAGCAGCTGCCCCAG 1800
Qy 3859 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAAGGATATGCTCTGA 3918
Db 1801 ACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGCAAGGATATGCTCTGA 1860
Qy 3919 CAGGAACAGTCTGTGGATGGACATGATCAGTGTAA--GGAAAGCAGCAGAGAGAGAGC- 3975
Db 1861 CAGGAACAGTCTGTGGATGGACATGATCAGTGTAAAGGTAAGGTAAGCAGCAGAGAGAGAGC 1920
Qy 3976 TCCGGCGCCCCAGCCCCCACTATCAGTGT-CCAGCGTGTGTTTCCCCAGAGCAGACAGC- 4031
Db 1921 TCCGGCGCCCCAGCCCCCACTATCAGTGTGTTTCCCCAGAGCAGACAGCAGCAGCAGCAGC 1980
Qy 4032 -----TCAGATATCACTGACACTCACCTGTCGCCCTGCCCCCTGG--CCAGAGGGTACTG 4082
Db 1981 GCTNACAGCATCANCACTGACACTNCACTCCCTTNGCCCTGCCCCCTGCCCCCANGAGGGTACTG 2040
Qy 4083 CCG-ACGGCACTTTGCACTCTG--ATGACCTCAAGCAGCTTTCATGCTGCTGCCCTCTGCGCA 4139
Db 2041 CCGNACGGCACTTTGCACTCTGATGACCTTCAAGCAGCTTTCATGCTGCTGCCCTCTGCGCA 2100
Qy 4140 GGGCAGGGCAGGGCAGTGAACACT-----GTAGGAGCATAGCAGCAGCAGAGATGG 4189
Db 2101 GCAGGNCAGGNCAGGNCAGTGCANCTGTAGGAGCATANGCAANGCCAGAGATGG 2160
Qy 4190 GGTGAA--GGGACACAGTCTGTAGCTGTCA-CATGATGTGACT-CCTCAAACTCTT- 4244
Db 2161 GGTGNAAGGGANCACTGTGAGCTGTCCANCACTGATGTGACTNCTCAAACTCTTN 2220
Qy 4245 --CCAGATTTCTTAAGAATAGCACCCCC--TTCCCCATTTGCCCTTCCAGCTTGTCTTCT 4300
Db 2221 NCCAGNATTTCTTAAGAATAGCANCCCCCTTNCCTTNCCTTNCCTTNCCTTNCCTTNCCTT 2280
Qy 4301 CCCAGGGGAGCTA-CTCAGGACTCACGTAGCATTAATCAATCAGTGTG-AATCGTCAGGGGG 4358
Db 2281 CCCAGGGGAGCTANTCAGAGCTCACGTAGCATTAATCAATCAGTGTGNAATCGTCAGGGGG 2340
Qy 4359 TGTCTGTAGCCTCAACCTCTGTGGGAGGGGAGCGCCGAGACTCCGTGGGAGAGCTCAT 4418
Db 2341 TGTCTGTAGCCTCAACCTCTGTGGGAGGGGAGCGCCGAGACTCCGTGGGAGAGAGCTCAT 2400
Qy 4419 TCCACATCTTGCAGAGAGCCTTT--GTCCAGCTGTCCACATGATGATGATGATGCTCCC 4477
Db 2401 TCCACATCTTGCAGAGAGCCTTTTNGTCCAGCTGTCCACATGATGATGATGATGCTCCC 2460
Qy 4478 GGGGAGAGAGCCCCGGCCCCCAGACATAAAGACTCAGCCTTGGTACTGCGAGTCTG 4537
Db 2461 GGGGAGAGAGCCCCGGCCCCCAGACATAAAGAACTCAGCCTTGGTACTGCGAGTCTG 2520
Qy 4538 GGTGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGTATGACAAATGTAAAAAAA 4596
Db 2521 GGTGTGTAGAGAACTCTTTTGAAGCAATAAAGTTTGGGGTGTATGACAAATGTAAAAAAA 2579
```

RESULT -6

US-60-659-397-12129

```
; Sequence 12129, Application US/60659397
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CHANG, Sheng-Yung
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSE TO INTERPERON TREATMENT IN HEPATITIS C
; TITLE OF INVENTION: VIRUS-INFECTED SUBJECTS, METHODS OF DETECTION AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001470
; CURRENT APPLICATION NUMBER: US/60/659,397
; CURRENT FILING DATE: 2005-03-09
; NUMBER OF SEQ ID NOS: 47859
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12129
; LENGTH: 65967
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-659-397-12129
```

Query Match 33.8%; Score 1554.2; DB 15; Length 65967;  
Best Local Similarity 97.7%; Pred. No. 0;  
Matches 1645; Conservative 5; Mismatches 23; Indels 10; Gaps 7;

```
Qy 2909 AGATCCAGCTGCAGCCTTCAGCTTGTGTACCAAAAGAGCGGCAGCCTGTTGCTACGACA 2968
Db 58244 AGATCCAGCTGCAGCCTTCAGCTTGTGTACCAAAAGAGCGGCAGCCTGTTGCTACGACA 58303
Qy 2969 TGAGAGTGCAGACTGGGCATCGACCTGCAGTGCACACTGGCCCTGTATGGCAGCTTCG 3028
Db 58304 TGAGAGTGCAGACTGGGCATCGACCTGCAGTGCACACTGGCCCTGTATGGCAGCTTCG 58363
Qy 3029 CTTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAAACAGGCCCTAACCCCTGCCCTCCAC 3088
Db 58364 CTTGGAGCTGGAGGGTCAAGCATGGCCAGCTGGAGAAACAGGCCCTAACCCCTGCCCTCCAC 58423
Qy 3089 CGCGGCTCCACACTGCGGAAAGAGCCTTCTCTGCTGGTGCAGATGCTGCTGCTGAAA 3148
Db 58424 CGCGGCTCCACACTGCGG--AAGCAGCCTTCTCTGCTGGTGCAGATGCTGCTGCTG-AA 58481
Qy 3149 ACACAGCTCAGCGTTTCCAGGGGATTTGCGAGCCCCCGCTCCAGTGGGAACACAGGG 3208
Db 58482 ACACAGCTCAGCGTTTCCAGGGATCTGCGAG--CCCCCGCTCAGCAGTGGGACACAGGG 58540
Qy 3209 CTTGCGAGCAGCAAGTGGGGGCAAGCAGAAATGCTCTCCAGGATTTTCAACACTGAGCCCT 3268
Db 58541 CTTGCGAGCAGCAAGTGGGGGCAAGCAGAAATGCTCTCCAGGATTTTCAACACTGAGCCCT 58600
Qy 3269 GCGCCACCTGCTGAAAAAACA--TCCGCCAGCTGAAGAGACAGAGAGAGATGCGAGAG 3327
Db 58601 GCGCCACCTGCTGAAAAAACA--TCCGCCAGCTGAAGAGACAG--AGGAGATGCGAGAG 58659
Qy 3328 TTACCTGGGGAACAAACAGGGATCTTTTCTGCCCCCTCTCCAGTCCAGTGGCCCTGA 3387
Db 58660 TTACCTGGGGAACAAACAGG---ATCTTCTCTGCCCCCTCTCCAGTGGCCCTGA 58715
Qy 3388 CCAGCTTGAATCAGTACCATTTTGTGGCAGACAGGGGAGAGCAGCTTTCAGACCTGGGTC 3447
Db 58716 CCAGCTTGAATCAGTACCATTTTGTGGCAGACAGGGGAGAGCAGCTTTCAGACCTGGGTC 58775
Qy 3448 AGAAGGGGTGGGCGAGCCCTTCCGCCCTTCCAGCTGCTGTGAGAGTGTCAAGT 3507
Db 58776 AGAAGGGGTGGGCGAGCCCTTCCGCCCTTCCAGCTGCTGTGAGAGTGTCAAGT 58835
Qy 3508 GTGTAGGGGCCCCAACTCAGGTTTCAGTGCAGAACAGGTCAGCAGGATGTCGCCGCCCTGA 3567
Db 58836 GTGTAGGGGCCCCAACTCAGGTTTCAGTGCAGAACAGGTCAGCAGGATGTCGCCGCCCTGA 58895
Qy 3568 GGTAAAGGGGCGCTCTTAAACCCCTTGGCTGGCTCACCTGGCCAGCTCACCCCTTTTGG 3627
Db 58896 GGTAAAGGGGCGCTCTTAAACCCCTTGGCTGGCTCACCTGGCCAGCTCACCCCTTTTGG 58955
Qy 3628 GTGTAGGGGAAAAAAGATGCTGACCTTGGGAAGGCTCCCTGGTAGAATAACACCACTTTT 3687
```



Db 58956 GTGTAGGGGAAAGAAATGCTGACCCCTGGGAAGGCTCCCTGGTAGAATACACACACTTT 59015  
QY 3688 TCAGGTTGTTGCAACACACAGTCTCGATTGACCTCTGGTTTACGCAAGAGCAACAAAGAGG 3747  
Db 59016 TCAGGTTGTTGCAACACACAGTCTCGATTGACCTCTGGTTTACGCAAGAGCAACAAAGAGG 59075  
QY 3748 TGTGTAAGTGAAGTGTCTCAGTCCCAAGACATGTGCCCCCTTTGCTGTGCTACCACT 3807  
Db 59076 TGTGTAAGTGAAGTGTCTCAGTCCCAAGACATGTGCCCCCTTTGCTGTGCTACCACT 59135  
QY 3808 CTTCCCAAGACAGAGAGGCCCCGAGCCCTTCAGGCCACGACTGCCCCAGACTCGCTGG 3867  
Db 59136 CTTCCCAAGACAGAGAGGCCCCGAGCCCTTCAGGCCACGACTGCCCCAGACTCGCTGG 59195  
QY 3868 CACTCAGTTCCCTCATCTGTAAAGTGAAGGGTGTATGCAAGGATATGCTGTGACAGGAACAG 3927  
Db 59196 CACTCAGTTCCCTCATCTGTAAAGTGAAGGGTGTATGCAAGGATATGCTGTGACAGGAACAG 59255  
QY 3928 TCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAGAGAGAGAGC-TCGCGGCCCC 3986  
Db 59256 TCTGTGGATGGACATGATCAGTGTCTAAGGAAAGCAGAGAGAGAGCCTCCGCGGCCCC 59315  
QY 3987 AGCCCACTATCAGTGTCCAGGCTGCTGGTTCCCAAGACAGACTCAGCATCACACTGA 4046  
Db 59316 AGCCCACTATCAGTGTCCAGGCTGCTGGTTCCCAAGACAGACTCAGCATCACACTGA 59375  
QY 4047 CACTCACCTGCTGCCCCCTGGCCAGAGGGTACTGCCAGCGCACTTTGCACTCTGATG 4106  
Db 59376 CACTCACCTGCTGCCCCCTGGCCAGAGGGTACTGCCAGCGCACTTTGCACTCTGATG 59435  
QY 4107 ACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGCGCAGGCGAGTGCACCTGTAG 4166  
Db 59436 ACCTCAAGCACTTTTCATGGCTGCCCTCTGGCAGGCGCAGGCGAGTGCACCTGTAG 59495  
QY 4167 GAGCATAGCAGCAGAGAGTGGGTGAAGGACACAGTCTTGACCTGTCCACATGATG 4226  
Db 59496 GAGCATAGCAGCAGAGAGTGGGTGAAGGACACAGTCTTGACCTGTCCACATGATG 59555  
QY 4227 TGACTCCTCAAACTCTTCCAGATTCTCTAAGATAGCACCCCTTCCCAATTGCCCA 4286  
Db 59556 TGACTCCTCAAACTCTTCCAGATTCTCTAAGATAGCACCCCTTCCCAATTGCCCA 59615  
QY 4287 GCTTAGCTCTTCTCCAGGGAGTACTCAGGACTCACTAGATTAATAACAGTGTGA 4346  
Db 59616 GCTTAGCTCTTCTCCAGGGAGTACTCAGGACTCACTAGATTAATAACAGTGTGA 59675  
QY 4347 ATCGTCAGGGGTGTCTGTAGCTCAACCTCCTGGGCGAGGGGAGACTCCGTG 4406  
Db 59676 ATCGTCAGGGGTGTCTGTAGCTCAACCTCCTGGGCGAGGGGAGACTCCGTG 59735  
QY 4407 GGAGAAGCTCATTTCCACATCTTGCAAGACAGACCTTTGTCCAGCTGTCCACATTGATC 4466  
Db 59736 GGAGAAGCTCATTTCCACATCTTGCAAGACAGACCTTTGTCCAGCTGTCCACATTGATC 59795  
QY 4467 AGACTGCTCCGGGAGAGAGCCCCCGCCCCCAGCACATAAAGAACTCCAGCCTTGGTAC 4526  
Db 59796 AGACTGCTCCGGGAGAGAGCCCCCGCCCCCAGCACATAAAGAACTCCAGCCTTGGTAC 59855  
QY 4527 TGCAGAGTCTGGGTGTAGAGAACTTTGTAAAGCAATAAAGTTTGGGGTGTAGCAAAAT 4586  
Db 59856 TGCAGAGTCTGGGTGTAGAGAACTTTGTAAAGCAATAAAGTTTGGGGTGTAGCAAAAT 59915  
QY 4587 GTT 4589  
Db 59916 GTT 59918

RESULT 7

US-10-450-763-15295/c  
; Sequence 15295, Application US/10450763  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 15295  
; LENGTH: 2760  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (1282)..(1566)  
; OTHER INFORMATION: 73% homologous to Mus musculus Pro-Pol-dUTPase  
; OTHER INFORMATION: polyprotein, accession number Y12713, Smith-Waterman Score=381.  
US-10-450-763-15295

Query Match 32.1%; Score 1474.4; DB 8; Length 2760;  
Best Local Similarity 73.4%; Pred. No. 0;  
Matches 2370; Conservative 0; Mismatches 6; Indels 852; Gaps 4;

QY 232 ATGCGAGTGAAGAAATGGCCTGCCAGGTGCCCTCGCTCAGCAGTGGGCGAGCAGAG 291  
Db 2760 ATGGCAGTGAAGAAATGGCCTGCCAGGTGCCCTCGCTCAGCAGTGGGCGAGCAGAG 2701  
QY 292 GAACTCCCAAGCCAAAGGAGAGACGCCGCCACTTGGGGAAGAAACAGAGCTCCGTCTTAC 351  
Db 2700 GAACTCCCAAGCCAAAGGAGAGACGCCGCCACTTGGGGAAGAAACAGAGCTCCGTCTTAC 2641  
QY 352 AAGCTTAGGCCCTGTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 411  
Db 2640 AAGCTTAGGCCCTGTGAGAGAGAGCCCTGTGTTCTGCGGAAAGTGGAGATCCTGAATGAC 2581  
QY 412 GTGATTACCAAGGCACAGCAAGAGGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 471  
Db 2580 GTGATTACCAAGGCACAGCAAGAGGCTCCGAGGCGAGGCCAGCTGCCATCTCTATC 2521  
QY 472 ATCCGCCAGGCTGAGT----- 487  
Db 2520 ATCCGCCAGGCTGAGTGTAGGGGCTAGGGCTTTGTTGGGCACGGGCGAGGCCCTTACC 2461  
QY 488 ----- 487  
Db 2460 ATCTTCCACCTAGAGGTGGCCCCCAGCCACAGGAATGAGCACCGTCTCTTGAGCCTGCTG 2401  
QY 488 -----GTGAGAAATAGCCAAGAGTTCAGGCCCAACC 516  
Db 2400 GGCCAGTCTGGGGGGGAGGCTGCTGAGGGTGAAGATAGCCAAGAGTTCAGGCCCAACC 2341  
QY 517 TTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGATACAGCCAGTCCGAGAGTCTTGAT 576  
Db 2340 TTTTCAGAACGCAATTTTCATCGCTGGGTCCAAACAGATACAGCCAGTCCGAGAGTCTTGAT 2281  
QY 577 CAGATCCCAACAATGTGGCCCATCTGACAGAGGCAAAATGGCCCGTGTGTGTGGAG 636  
Db 2280 CAGATCCCAACAATGTGGCCCATCTGACAGAGGCAAAATGGCCCGTGTGTGTGGAG 2221  
QY 637 GGAAAGCGTCCAGCAAAAGCCCGGAAGAAACGGAAGAGAGAGCTCAAAAGTCCCTGGCT 696  
Db 2220 GGAAAGCGTCCAGCAAAAGCCCGGAAGAAACGGAAGAGAGAGCTCAAAAGTCCCTGGCT 2161  
QY 697 CATCAGAGTGGCTTTGGCCAAACCCCTTCCAGGACCCCTGAGCAGGAGAGTGCACC 756  
Db 2160 CATCAGAGTGGCTTTGGCCAAACCCCTTCCAGGACCCCTGAGCAGGAGAGTGCACC 2101  
QY 757 ATCCAGTGCAGGAGATGAGTCTCCACTCGGCCCCCATATGTTAGNAACACACCCGAG 816  
Db 2100 ATCCAGTG----- 2092





Db 348 AGCCAGGCGCAGGCTCGAAGCTCCAGCTGGAACATGGTGTGCGCCGCGGGCGGCCACC 289  
QY 2788 GACACCCCAAGCTATTTCATAGGTGTGAAGTCCAAATACAGTCTCTTAATGGTGAACAC 2847  
Db 289 GACACCCCAAGCTATTTCATAGGTGTGAAGTCCAAATACAGTCTCTTAATGGTGAACAC 229  
QY 2848 CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATACGACAGC 2907  
Db 228 CTGCACATCCGGGAGTTCCACCGGGTCAAAGTGGGAGACATCGCCACTGGCATACGACAGC 169  
QY 2908 CAGATCCCAAGCTGCGACCTTCAGCTTGGTCAACAAAGACGGGACGCTGTTGCTTACGAC 2967  
Db 168 CAGATCCCAAGCTGCGACCTTCAGCTTGGTCAACAAAGACGGGACGCTGTTGCTTACGAC 109  
QY 2968 ATGGAGGTGCCAGACTCGGGCATACACCTGAGTGCACACTGGCCCTCGATGGGAGCTTC 3027  
Db 108 ATGGAGGTGCCAGACTCGGGCATACACCTGAGTGCACACTGGCCCTCGATGGGAGCTTC 49  
QY 3028 GCCTGGAGCTGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTAA 3075  
Db 48 GCCTGGAGCTGAGGGTCAAGCATGGCCAGCTGGAGAACAGGCCCTAA 1

RESULT 8  
US-11-060-756-560  
; Sequence 560, Application US/11060756  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; TITLE OF INVENTION: Mounts, William Martin  
; FILE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug  
; FILE OF INVENTION: Target Genes  
; FILE REFERENCE: AM101083 (031896-042000)  
; CURRENT APPLICATION NUMBER: US/11/060,756  
; CURRENT FILING DATE: 2005-02-18  
; NUMBER OF SEQ ID NOS: 303284  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 560  
; LENGTH: 1400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-060-756-560

Query Match 29.3%; Score 1345.6; DB 13; Length 1400;  
Best Local Similarity 98.6%; Pred. No. 5.3e-293;  
Matches 1384; Conservative 5; Mismatches 9; Indels 6; Gaps 3;  
QY 3195 AGTGGAAACCAAGGCTCGCAGCAGCAAGGTGGGGCAAGAGATGCTCCAGGATTT 3254  
Db 1 AGTGGAAACCAAGGCTCGCAGCAGCAAGGTGGGGCAAGAGATGCTCCAGGATTT 60  
QY 3255 CACACCTGAGCCCTGCCCCACCTGCTGAATAACA-TCCGCCAGCTGAAGAGACAGAAG 3313  
Db 61 CACACCTGAGCCCTGCCCCACCTGCTGAATAACA-TCCGCCAGCTGAAGAGACAGAAG 120  
QY 3314 GAGGATGCGAGGAGTTACTCTGGGAAACAAACAGGGATCTTTTCTGCCCTGCTCCAG 3373  
Db 121 GAGGATGCGAGGAGTTACTCTGGGAAACAAACAGG---ATCTTCTGCTGCTCCAG 176  
QY 3374 TCGAGTTGGCTGACCCGCTGGATCAGTGACCATTTGTCGACAGAGGGAGAGCAGC 3433  
Db 177 TCGAGTTGGCTGACCCGCTGGATCAGTGACCATTTGTCGACAGAGGGAGAGCAGC 236  
QY 3434 TTCCAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCTCCAGGCTGCTG 3493  
Db 237 TTCCAGCTGGGTGAGAGGGGTGGCGAGCCCTTCGCGCCCTCACCTCCAGGCTGCTG 296  
QY 3494 TGAGAGTGTCAAGTGTGAAGGGGCCAAACTCAGGTTTCAGTGCAGAACCCAGCTCAGCAGG 3553  
Db 297 TGAGAGTGTCAAGTGTGAAGGGGCCAAACTCAGGTTTCAGTGCAGAACCCAGCTCAGCAGG 356  
QY 3554 TATGCCGCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGGCTGGCCCTCACCTGGGCCAG 3613

Db 357 TATGCCGCCCGTAGGTTAAGGGGCCCTCTAAACCCCTTGGCTGGCCCTCACCTGGGCCAG 416  
QY 3614 CTCACCCCTTTTGGGTGTAGGGGAAAAAATGCTGACCCCTGGGAGAGCTCCCTGGTAGA 3673  
Db 417 CTCACCCCTTTTGGGTGTAGGGGAAAAAATGCTGACCCCTGGGAGAGCTCCCTGGTAGA 476  
QY 3674 ATACACACACACTTTTTCAGGTTGTTCACACACAGAGTCTGAGTTGACCTCTGGTTTCAGCCA 3733  
Db 477 ATACACACACACTTTTTCAGGTTGTTCACACACAGAGTCTGAGTTGACCTCTGGTTTCAGCCA 536  
QY 3734 AGGACCAAGAGGTGTGTAAGTGAAGTGTTCCTCAGTCCCAGACATGTGCCCTTTGTC 3793  
Db 537 AGGACCAAGAGGTGTGTAAGTGAAGTGTTCCTCAGTCCCAGACATGTGCCCTTTGTC 596  
QY 3794 TGCTGGCTTACACACTTCTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCACGACCTGTC 3853  
Db 597 TGCTGGCTTACACACTTCTCCAGAGCAGAGCCCGGAGCCCTTCAGGCCACGACCTGTC 656  
QY 3854 CCCAGACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGACGATATG 3913  
Db 657 CCCAGACTCGCTGGCAGCTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATGACGATATG 716  
QY 3914 CCTGACAGGAAACAGTCTGTGGATGACATGATGATGATGATGATGATGATGATGATGATG 3973  
Db 717 CCTGACAGGAAACAGTCTGTGGATGACATGATGATGATGATGATGATGATGATGATGATG 776  
QY 3974 CG-TCCGGCGCCAGCCCGCAGCTCAGTGTCCAGCTGCTGGTTCCTCCAGAGCAGCT 4032  
Db 777 CGCTCCGGCGCCAGCCCGCAGCTCAGTGTCCAGCTGCTGGTTCCTCCAGAGCAGCT 836  
QY 4033 CAGCATCACACTGACACTCACCTCGCTGCCCTGGCGAGAGGTACTCGCGACGGCAC 4092  
Db 837 CAGCATCACACTGACACTCACCTCGCTGCCCTGGCGAGAGGTACTCGCGACGGCAC 896  
QY 4093 TTTGACCTCTGATGACCTCAAAGCAGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4152  
Db 897 TTTGACCTCTGATGACCTCAAAGCAGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956  
QY 4153 CAGTGAACCTGTAGGAGATAGCAGCCAGAGATGGGGTGAAGGGACACAGTCTTTGAGC 4212  
Db 957 CAGTGAACCTGTAGGAGATAGCAGCCAGAGATGGGGTGAAGGGACACAGTCTTTGAGC 1016  
QY 4213 TGTCCACATGCTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCT 4272  
Db 1017 TGTCCACATGCTGACTCTCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCT 1076  
QY 4273 TCCCATTTGCCAGCTTAGCTCTTCTCCAGGGAGCTACTCAGGACTCACGTAGCAT 4332  
Db 1077 TCCCATTTGCCAGCTTAGCTCTTCTCCAGGGAGCTACTCAGGACTCACGTAGCAT 1136  
QY 4333 TAAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCTCAACCTCTGGGGCAGGGGAC 4392  
Db 1137 TAAATCAGCTGTGAATCGTCAAGGGGTGTCTGCTAGCTCAACCTCTGGGGCAGGGGAC 1196  
QY 4393 GCCGAGACTCCGTTGGGAGAGCTCATTTCCACACTTTTGGCCAAAGACAGCTTTTTCAGCT 4452  
Db 1197 GCCGAGACTCCGTTGGGAGAGCTCATTTCCACACTTTTGGCCAAAGACAGCTTTTTCAGCT 1256  
QY 4453 GTCCACATTTAGTCAAGTGTCTCCCGGGAGAGAGCCCGGCCCCAGCACAATAAAGAAC 4512  
Db 1257 GTCCACATTTAGTCAAGTGTCTCCCGGGAGAGAGCCCGGCCCCAGCACAATAAAGAAC 1316  
QY 4513 TGCAGCTTTGGTACTGAGAGCTGTGGGTGTGAGAGAACTCTTTCTAGCAATAAAGTTG 4572  
Db 1317 TGCAGCTTTGGTACTGAGAGCTGTGGGTGTGAGAGAACTCTTTCTAGCAATAAAGTTG 1376  
QY 4573 GGGTGTGACAAATGTTAAAAAAA 4596  
Db 1377 GGGTGTGACAAATGTTAAAAAAA 1400

Sequence 4832, Application US/11060756  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William Martin  
; TITLE OF INVENTION: Nucleic Acid Arrays for Monitoring Expression Profiles of Drug  
; TARGET GENES:  
; FILE REFERENCE: AM101083 (031896-042000)  
; CURRENT APPLICATION NUMBER: US/11/060,756  
; CURRENT FILING DATE: 2005-02-18  
; NUMBER OF SEQ ID NOS: 303284  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4832  
; LENGTH: 1400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-060-756-4832

Query Match 29.3%; Score 1345.6; DB 13; Length 1400;  
Best Local Similarity 98.6%; Pred. No. 5.3e-293;  
Matches 1384; Conservative 5; Mismatches 9; Indels 6; Gaps 3;  
3195 AGTGGGAACACAGGCGCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATCGCTCCCAAGGATTT 3254  
1 AGTGGGAACACAGGCGCTCGCAGCAGCAAGGTGGGGGCAAGCAGAAATCGCTCCCAAGGATTT 60  
3255 CACACCTGAGCCCTGCCCCACCTGCTGAAAAACA-TCCGCCACGTGAAGACAGAAAG 3313  
61 CACACCTGAGCCCTGCCCCACCTGCTGAAAAACA-TCCGCCACGTGAAGACAGAAAG 120  
3314 GAGATGGCAGGAGTACCTGGGGAAACAAACAGGAGATCTTTTCTGCCCTGCTCCAG 3373  
121 GAGATGGCAGGAGTACCTGGGGAAACAAACAGGAGATCTTTTCTGCCCTGCTCCAG 176  
3374 TCCAGTTGGCTGACCCGCTGGATCAGTACCAATTTTGGCAGACAGGGAGAGCAGC 3433  
177 TCGAGTTGGCTGACCCGCTGGATCAGTACCAATTTTGGCAGACAGGGAGAGCAGC 236  
3434 TCCAGCTGGCTGAGAGGGGTGGGGAGCCCTTCGGCCCTCACCCTCCAGGCTGCTG 3493  
237 TCCAGCTGGCTGAGAGGGGTGGGGAGCCCTTCGGCCCTCACCCTCCAGGCTGCTG 296  
3494 TGAGAGTGTCAAGTGTAGAGGGGCCAAATCTCAGGTTTCAAGTGTCAAGTGTCAAGTGT 3553  
297 TGAGAGTGTCAAGTGTAGAGGGGCCAAATCTCAGGTTTCAAGTGTCAAGTGTCAAGTGT 356  
3554 TATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCCTCACCTGGCCAG 3613  
357 TATGCCCGCCGTAGGTTAAGGGGGCCCTCTAAACCCCTTGCCTGGCCCTCACCTGGCCAG 416  
3614 CTACACCCCTTTTGGGTGTAGGGGAAAAAGAAATGCTGACCCCTGGGAAGGCTCCCTGGTAGA 3673  
417 CTACACCCCTTTTGGGTGTAGGGGAAAAAGAAATGCTGACCCCTGGGAAGGCTCCCTGGTAGA 476  
3674 ATACACACACTTTTTCAGGTTTGTGCAACACAGCTCTGAGTTGACCTCTGTTTCAGCCA 3733  
477 ATACACACACTTTTTCAGGTTTGTGCAACACAGCTCTGAGTTGACCTCTGTTTCAGCCA 536  
3734 AGGACCAAGAGTGTGTAAGTCAAGTGTCTCAGTCCCAAGACATGTGCCCTTTGC 3793  
537 AGGACCAAGAGTGTGTAAGTCAAGTGTCTCAGTCCCAAGACATGTGCCCTTTGC 596  
3794 TGCTGGCTTACCACTTTTCCCAAGACAGCAGGCCCCCGAGCCCTTCAGGGCCACAGCTGC 3853  
597 TGCTGGCTTACCACTTTTCCCAAGACAGCAGGCCCCCGAGCCCTTCAGGGCCACAGCTGC 656  
3854 CCCAGACTCGCTGCACCTAGTTCCCTCATCTGTAAAGGTGAAGGGGTGATGACGAGATATG 3913  
657 CCCAGACTCGCTGCACCTAGTTCCCTCATCTGTAAAGGTGAAGGGGTGATGACGAGATATG 716  
3914 CCTGACAGGACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAGCAGCAGAGAGAGA 3973  
717 CCTGACAGGACAGTCTGTGGATGGACATGATCAGTCTTAAGGAAGCAGCAGAGAGAGA 776

Qy 3974 CG-TCCGGCGCCCCAGCCCCACATATCAGTGTCCAGCGTGTGTTCCCAAGACAGCAGCT 4032  
Db |||||  
777 CGCTCCGGCGCCCCAGCCCCACATATCAGTGTCCAGCGTGTGTTCCCAAGACAGCAGCT 836  
Qy 4033 CAGCATCACACTGACACTCACCCTGCCCTCGCCCTGCGCAGAGAGGTACTGCCGACGGCAC 4092  
Db |||||  
837 CAGCATCACACTGACACTCACCCTGCCCTCGCCCTGCGCAGAGAGGTACTGCCGACGGCAC 896  
Qy 4093 TTTGCACTCTGATGACCTCAAAGCATTTCATGGCTGCCCTCTGGCAGGGCAGGGCAGGG 4152  
Db |||||  
897 TTTGCACTCTGATGACCTCAAAGCATTTCATGGCTGCCCTCTGGCAGGGCAGGGCAGGG 956  
Qy 4153 CAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACAGTCTTGAGC 4212  
Db |||||  
957 CAGTGACACTGTAGGAGCATAGCAAGCCAGAGATGGGGTGAAGGGACACAGTCTTGAGC 1016  
Qy 4213 TGTCCCATGTCATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCT 4272  
Db |||||  
1017 TGTCCCATGTCATGTGACTCTCAAACTCTTCCAGATTTCTCTAAGAAATAGCACCCCT 1076  
Qy 4273 TCCCATTTGCCCGCAGCTTTCCTCCAGGGGAGCTACTCAGGACTCACGTAGCAT 4332  
Db |||||  
1077 TCCCATTTGCCCGCAGCTTTCCTCCAGGGGAGCTACTCAGGACTCACGTAGCAT 1136  
Qy 4333 TAAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAAACCTCTCTGGGGCAGGGGAC 4392  
Db |||||  
1137 TAAATCAGCTGTGAATCGTCAGGGGGTGTCTGTAGCCTCAAACCTCTCTGGGGCAGGGGAC 1196  
Qy 4393 GCCGAGACTCCGTGGGAGAGCTATTCACACATCTTCCCAAGACAGCCTTTGTCCAGCT 4452  
Db |||||  
1197 GCCGAGACTCCGTGGGAGAGCTATTCACACATCTTCCCAAGACAGCCTTTGTCCAGCT 1256  
Qy 4453 GTCCACATTCAGTCAGACTGCTCCCGGGAGAGAGCCCCCGCCCGCCAGCAGCATAAAGAAC 4512  
Db |||||  
1257 GTCCACATTCAGTCAGACTGCTCCCGGGAGAGAGCCCCCGCCCGCCAGCAGCATAAAGAAC 1316  
Qy 4513 TGAGCCTTGGTACTGACAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTG 4572  
Db |||||  
1317 TGAGCCTTGGTACTGACAGAGTCTGGGTGTAGAGAACTCTTTGTAAGCAATAAAGTTTG 1376  
Qy 4573 GGCTGATGACAAATGTTAAAAA 4596  
Db |||||  
1377 GGCTGATGACAAATGTTAAAAA 1400

RESULT 10  
US-60-544-32424  
; Sequence 32424, Application US/60680544  
; GENERAL INFORMATION:  
; APPLICANT: Cooper, Matthew  
; APPLICANT: Kinch, Deborah  
; APPLICANT: Rosenberg, Michael  
; APPLICANT: Subramaniam, S. Sai  
; APPLICANT: Szak, Suzanne  
; APPLICANT: Li, Huo  
; APPLICANT: Bandaru, Raj  
; APPLICANT: Derbel, Maher  
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary  
; TO INVENTION: Fragments of, Cynomolgus Monkey Genes and the Use Thereof  
; FILE REFERENCE: 21590290000  
; CURRENT APPLICATION NUMBER: US/60/680,544  
; CURRENT FILING DATE: 2005-05-13  
; NUMBER OF SEQ ID NOS: 48714  
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0  
; SEQ ID NO 32424  
; LENGTH: 3658  
; TYPE: DNA  
; ORGANISM: Macaca Mulatta  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)....(3658)  
; OTHER INFORMATION: n = A,T,C or G  
US-60-544-32424

Query Match		14.7%;	Score 676.2;	DB 15;	Length 3658;
Best Local Similarity		22.6%;	Pred. No. 4.5e-142;		
Matches 832;		Conservative 0;	Mismatches 2824;	Indels 30;	Gaps 10;
Qy	906	ACTCACAAACTGATCAGCCCTTGCATGTGTGAACCAACGCTGTGGAACTCACCACCC	965		
Db	1	ACTCCACAACTGATCAGCCCTTGCATGTGTCAACCAACGCTGTGGAACTCACCACCA-NX	59		
Qy	966	CCAGAGCGGAGGCCCTGCCCTGCGCCACGACCCCTTCCCTATATAGAGACTGCCTCA	1025		
Db	60	NN	119		
Qy	1026	TCCCTTCCCATTCACCTCTCCAGCCTGGAACTCACCCTCTGGAGTCCTCTGGG	1085		
Db	120	TCCCTTCCCATTCACCTCTCCAGCCTGGAACTCACCCTCTAGAGTCCTTCTGGG	179		
Qy	1086	CAAACTGGCTGTGA-GACAGCCAGAAACCTTGCCTGACCCACACCTGAGCAAACTGG	1144		
Db	180	CAAACTGGCTGTGAANN	239		
Qy	1145	CCTGTGTAGACAGTCAAAGCCCTGCTGGCCACACCTGGAGCCAGCTGCCTGTCTC	1204		
Db	240	NN	299		
Qy	1205	GTGGTGCCCATGAGAAGTTTCTGTGGAGGAATACCTAGTCATGCTCTGCAAGGCAGG	1264		
Db	300	NN	359		
Qy	1265	TGAGCTCAAGCCAGGCCACACGCTGACGAGCTGGCCAAAGACCTGGGCAGCAGCGGCT	1324		
Db	360	NN	419		
Qy	1325	CCAGATCCGGGAGCCGACCCCAAACTGAGGA CAACGAGGCTCTCTGTCACTGAGA	1384		
Db	420	NN	479		
Qy	1385	AACTCAAGCCAGTGTATGATACCGAGAAGTCCACTGGGCCACGACCACTCC	1444		
Db	480	NN	539		
Qy	1445	GCTGGGCAGAGGCTCTTCGGAGAGTGCA CAGGATGGAGACAAGCAGACTGCTTCC	1504		
Db	540	NN	599		
Qy	1505	AGTGGCTGTCAAAGGTGCGCTGGAAGTATTTCTGGGCAGAGAGCTGATGGCATGTG	1564		
Db	600	NN	659		
Qy	1565	CAGGATTGACCTCACCCGAAATTGTCCTTTGTATGGAGCTGTGAGAGAAGGCTTGGG	1624		
Db	660	NN	719		
Qy	1625	TCAACATCTTCATGGAGCTGTGGAAGTGGCTCCCTGGCCAGCTGTGTCGAAGCAGG	1684		
Db	720	NN	779		
Qy	1685	GCTGTCTCCAGAGACCGGCCCTGTACTACTGGGCCAGGCCCTGGAGGCTGTGGAAT	1744		
Db	780	NN	839		
Qy	1745	ACCTCACTCAGAAAGGATTCTGCATGGGAGCTCAAAGCTGACAACTGCTCTGTCCA	1804		
Db	840	NN	899		
Qy	1805	CGATGGAGCAACGACCCCTCTGTGACTTTGGCCATGCTGTGTCTTCAACCTGATG	1864		
Db	900	NN	959		
Qy	1865	GCCTGGGAAAGTCTTGTCTCAGGGGACTATCTCCCTGGCAGACGACCATGGCTC	1924		
Db	960	NN	1019		

Qy	1925	CGGAGGTGCTGCTGGCAGGAGCTCGGAGCCCAAGGTGATGCTGTGGAGCAGCTGCTGTA	1984
Db	1020	NN	1079
Qy	1985	TGATCTGCACATGCTCAACGGCTGCCACCCCTGGACTCAGTTCTTCGAGGGCCGCTCT	2044
Db	1080	NN	1139
Qy	2045	GCCTCAAGATTGCGCAGCAGCCTCGCTGTGAGGGAGATCCACCCCTCTGCGCCCTC	2104
Db	1140	NN	1199
Qy	2105	TCACAGCCAGGCCATCCAGAGGGCTGAGGAAGAGCCCATCACCGCTGTCTGCAG	2164
Db	1200	NN	1259
Qy	2165	CGGAGCTGGAGGGAAGGTGAACCGGCCTACAGCAAGTGGGAGTCTGAAGAGCCCTT	2224
Db	1260	NN	1319
Qy	2225	GGAGGGAGAAATAAAGAA CCAAGACATCCACCGCCAAATCAAGCCAAATTACCA	2284
Db	1320	NN	1379
Qy	2285	CCCTCCATGCCAGCCGAGAGAGCTTTCGCCAAGGGCCCGAGGCCCGCAGCTGAGG	2344
Db	1380	NN	1439
Qy	2345	AGACAAAGGAGAGCCCTTAAGCTCCAGCCTCTCTCCACAGAGCCCGCCAGAGCCAA	2404
Db	1440	NN	1499
Qy	2405	ACAACTCTCTCCCTTGACTTTTGAGCAAGAGGAGTCTGGGATGTGGAAACCTTTAC	2464
Db	1500	NN	1559
Qy	2465	TGCTCTCTCGAGCAGCCCTGCCAGAACCCAGCTCACAGAGCGGAAAGCAACCG	2524
Db	1560	TGCTCTCTCGAGCAGCCCTGCCAGAACCCAGCTCACAGAGCGGAAAGCAACCT	1619
Qy	2525	TCCGGAGCAGAACTGCAGCAGCTGGAAATATTTCTCTCAACAGCCTGTCTCCAGC	2584
Db	1620	TCCGGAGCAGAACTGCAGCAGCTGGAAATATNNNNNNNNNNNNNNNNNNNNNNNN	1679
Qy	2585	CATTTTCTCGAGAGCAGGAGCAAAATCTCTCTGCTGCTCAGCATCGACAGCCTCTCC	2644
Db	1680	NN	1739
Qy	2645	TGTCGGATGACAGTGAGAGAAACCATCAAAGGCTCTCAAAGCTCGCGGACACCTGA	2704
Db	1740	NN	1799
Qy	2705	GCTCAGGCGTACACTCTCGAGCAGCCAGGCGGAGCTCGAAGCTCCAGCTGGAACATGG	2764
Db	1800	NN	1859
Qy	2765	TGCTGGCCGGGGGGCCACCGACATCCGGGAGTTCCA CCGGCTCAAGTGGGAG	2824
Db	1860	NN	1919
Qy	2825	TACAGTCTCTTAATGGTGAA CACCTGCACATCCGGGAGTTCCA CCGGCTCAAGTGGGAG	2884
Db	1920	NN	1979
Qy	2885	ACATCGCCACTGGCATCAGCAGCCAGATCCAGCTGCAGCCTTCAGCTTGGTCAACAAAG	2944
Db	1980	NN	2039
Qy	2945	ACGGGAGCCTGTTCGTGACATGGAGGTGCAGACTCGGGCATCGACCTGCAGTGCA	3004
Db	2040	NN	2099
Qy	3005	CAC'TGGCCCTGTATGGCAGCTTCGCTGGAGCTGGAGGTCAAGCATGGCAGCTGGAGA	3064



[illegible][illegible]

```
QY 3185 CCGGGCTCA CAGTGGGAAACAGGGCCCTCGCAGCAGCAAGGTGGGGCAAGCAGAAATGCCT 3244
Db 2280 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3245 CCAGGATTTCA CACTGAGCCCTGCCCCCTGCTGAAATAAACAATCCGCCAGGTGAAG 3304
Db 2340 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3305 AGACAGAAGGAGGATGCAGGAGTTACCTGGGGAACAAACAGGGATCTTTTCTGCC 3364
Db 2400 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3365 CTGCTCCAGTTCGAGTTGGCCCTGACCCGCTTGGATCAGTGACCAATTTGTTGGCAGACAGG 3424
Db 2460 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3425 GAGAGCAGCTTCAGCTCGGTGAGAAGGGTGGCGAGCCCTTCGGCCCTCACCTCC 3484
Db 2520 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3485 AGGCTGCTGTGAGAGTGTCAAGTGTGTAAAGGGCCAAACTCAGGTTTCAGTGCAACACAG 3544
Db 2580 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3545 GTGAGCAGGTATGCCGCCCGCTAGGTTAAGGGGGCCCTCTAAACCCCTTCGCTGGCCTCA 3604
Db 2640 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3605 CTTGGCCAGCTCACCCCTTTGGGTGTAGGGGAAAGAAATGCTGACCTCGGGAAGGCTC 3664
Db 2700 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3665 CTTGGTAGAATACACACACTTTTTCAGGTTGTGCAACACAGGTCTCTGAGTGCACCTCTG 3724
Db 2760 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3725 GTTCAGCCAAAGCAAGGAGGTGTAAAGTGAAGTGGTTCTCAGTCCCGCAGACATGTG 3784
Db 2820 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3785 CCCTTTGCTGGCTACACTCTTCCCGAGAGCAGCGCCCGGAGCCCTTCAGGCC 3844
Db 2880 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3845 CAGCACTGCCAGACTCGCTGGCACTCAGTTCCCTCATCTGTAAAGGTGAAGGTGATG 3904
Db 2940 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
QY 3905 CAGGATATGCTGACAGGAACAGTCTGTGGATGGACATGATGCTAAAGGAAGCAGC 3964
Db 3000 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGA
QY 3965 AGAGAGAGAGCTCGGGGCCCGACCCCACTACAGTGTCCAGGTGCTGGTCCCCAGA 4024
Db 3060 CGC-----TCGGGCACCCAGCCCACTGTGAGCGTCCAGCGTCCAGCGTCCGGGCCCGCCCGA 3111
QY 4025 GCACAGCTCAGCATCAGCTGACACTCACCTGCCCTGCCCTGGCCGAGAGGCTACTGCC 4084
Db 3112 GCACAGC-----TCACAAAAACACTCACCCTGCCCTGCCCGCGGAGGCTACTGCT 3165
QY 4085 GAGGCACCTTTGCACTCTGATGACCTCAAAAGCACTTTTCATGGGTGCCCTCTGGCAGGCA 4144
Db 3166 GACAGCACTTTGCACTCTGATGACCTCAAGACACTTTCATGGGTGCCCTCT-----GGCA 3220
QY 4145 GGCAGGCGAGTGACACTGTAGGAGCATAGCAAGCAGGAGATGGGGTGAAGGACACAG 4204
Db 3221 GGCAGGCTCAGTGACTCTGTAGGAGCATAGCAGTCCAGGAGATGGGGTGAAGG--ACAG 3278
QY 4205 TCTTGAGCTGTCCACATGCACTGCACTCCTCAAACTCTTCCAGATTTCTCTAAGATAG 4264
Db 3279 TCTTGAGCTGTACATGCACTGCACTCCTCAAACTCTTCCAGATTTCTCT---AATAG 3335
```

```
QY 4265 CACCCCTTCCCATTTGCCCCAGCTTAGCCTCTTCTCCAGGGAGGAGCTACTCAGGACTCA 4324
Db 3336 CAACCACTTCCCATTTGCCCCAGCTTAGCCTCTTCTCCAGGGAGGAGCTACTCAGGACTCA 3395
QY 4325 CGTAGCATTAATAACAGCTGTGAATCGTCA--GGGGGTGCTCTGTAGCCTCAACCTCTCTGGG 4383
Db 3396 TGTAGCATTAAGTTCAGCTGTGAATTTGTGAGGGGGTGTGAGTAGCCTCAACCTCTCTGGG 3455
QY 4384 GCAGGGGACGCCGAGAGCTCCGTGGGAGAGCTCATTTCCACATCTTTGCCAAGACAGCCTT 4443
Db 3456 GCGGGGAGCGCCGAGACTCTCTGGGAGAGCTCATCCACGCTTGTG--NNNNACAGCCTT 3514
QY 4444 TGTCCAGCTGTCCACATTTGAGTGTGAGTGTCTCCGGGGAGAGAGCCCCGGCCCCCAGCAC 4503
Db 3515 TGTCCAGCTGTCCACAGCGAGCAGACTGTCTCCGGGGAGACAGCCCCGGCCCCCAGCAC 3574
QY 4504 ATAAAGAACTGCAGCCTTGTGACTGTGAGAGTCTGGGTGTGAGAGAACTCTTTGTAAGCAA 4563
Db 3575 ATAAAGAACTGCAGCGTTGTGACTGTC--AGTGTGGGTGTGAGAGAACTCTTTGTAAGCAA 3632
QY 4564 TAAAGTTTGGGGTGTGATGACAAATGTT 4589
Db 3633 TAAAGTTTGGGGTGTGATGACAGATGTT 3658
```

## RESULT 12

```
US-10-450-763-15296
; Sequence 15296, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 15296
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (157)..(468)
; OTHER INFORMATION: 68% homologous to Mus musculus Pro-Pol-dUTPase
; OTHER INFORMATION: polyprotein, accession number Y12713, Smith-Waterman Score=354.
US-10-450-763-15296
```

```
Query Match 13.7%; Score 627.6; DB 8; Length 722;
Best Local Similarity 94.8%; Pred. No. 2.6e-131;
Matches 681; Conservative 0; Mismatches 34; Indels 3; Gaps 3;
```

```
QY 2564 TCCTCAACAGCCTGTCCAGCCATTTCTCTGGAGGAGGAGGAGCAATTTCTCTCGTGCC 2623
Db 5 TTCTTAACAGCCTGTCCAGCCATTTCTCTGGAGGAGGAGGAGCAATTTCTCTCGTGCC 64
QY 2624 TCAGCATCGACAGCCTCTCCCTCTCGGATGACAGTGAGAGAGAAACCCATCAAGGCTCTC 2683
Db 65 TCAGCATCGACAGCCTCTCCCTCTCGGATGACAGTGAGAGAGAAACCCATCAAGGCTCTC 124
QY 2684 AAAGCTCGGGGACACCTTGAGTCTAGGCGGTACACTCTCGAGCAGCAGCCGAGGCTC 2743
Db 125 AAAGCTCGGGGACACCTTGAGTCTAGGCGGTACACTCTCGAGCAGCAGCCGAGGCTC 184
QY 2744 GAAGCTCCAGCTGGACATGGTCTGTGGCCCGGGGGCCCGACCGACACCCAGCTATT 2803
Db 185 GAAGCTCCAGCTGGACATGGTCTGTGGCCCGGGGGCCCGACCGACACCCAGCTATT 244
```



```
QY 2804 TCAATGGTGTGAAGTCCAAATACAGTCTCTTAATAGTTGGAACACCTGCACATCCGGAGT 2863
Db |||||||
QY 245 TCAATGGTGTGAAGTCCAAATACAGTCTCTTAATAGTTGGAACACCTGCACATCCGGAGT 304
Db |||||||
QY 2864 TCCACCGGTCAAAGTGGGAGACATCGGCATCTGGCATCAGCAGCCAGATCCCGAGTGCAG 2923
Db |||||||
QY 305 TCCACCGGTCAAAGTGGGAGACATCGGCATCTGGCATCAGCAGCCAGATCCCGAGTGCAG 364
Db |||||||
QY 2924 CTTTCAGCTTGTGTACCAAGAGCGGCAGCCCTGATGGAGCTTTCGCTACGACATGAGGTGCCAGCT 2983
Db |||||||
QY 365 CTTTCAGCTTGTGTACCAAGAGCGGCAGCCCTGATGGAGCTTTCGCTACGACATGAGGTGCCAGCT 424
Db |||||||
QY 2984 CGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGAGCTTTCGCTGGAGCTGGAGGG 3043
Db |||||||
QY 425 CGGGCATCGACCTGCAGTGCACACTGGCCCTGATGGAGCTTTCGCTGGAGCTGGAGGG 484
Db |||||||
QY 3044 TCAAGCATGGCCAGCTGGAGAACAGCCCTAAACCTGCCCCTCCACCGCCGGCTCCACACT 3103
Db |||||||
QY 485 TCAAGCATGGCCAGCTGGAGAACAGCCCTAAACCTGCCCCTCCACCGCCGGCTCCACACT 544
Db |||||||
QY 3104 GCGGAAAGCAGCTTCTCTGC-TGGTGCACGATGC-TGCCCTGAAACACAGGCTCAGC 3161
Db |||||||
QY 545 TGCCGGAAGCAGCTTCTCTGCTTGGTGCACGATGCTTGCCCTTGAACACAGAGGCTCAGC 604
Db |||||||
QY 3162 CGTTCCCGAGGGAT-TGCCAGGCCCCCGGCTCACAGTGGGAACAGGCGCTCGCAGCAGC 3220
Db |||||||
QY 605 CGTTCCCGAGGGATCTTGGCAGCCCCCGGCTCAGCAGTGGGACAGGCGCTCGCAGCAGC 664
Db |||||||
QY 3221 AAGTGGGGGCAAGCAGAAATGCTCCAGAGATTTCCACACCTGAGCCCTGCCGCCACCT 3278
Db |||||||
QY 665 AAGTGGGGGCAAGCAGAAAGGCTCCCAAGATTTACAACCTGGGCGCTTGCCGCCCT 722
Db |||||||

RESULT 13
US-60-680-544-10785/c
; Sequence 10785, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE REFERENCE: 21590290000
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 10785
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; NAME/KEY: misc_feature
; LOCATION: (1)...(761)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-544-10785

Query Match 9.6%; Score 440.6; DB 15; Length 761;
Best Local Similarity 72.6%; Pred. No. 3.6e-89;
Matches 566; Conservative 0; Mismatches 192; Indels 22; Gaps 8;

QY 3777 GACATGTGCCCTTTGCTGTGGTACCACTCTTCCCGCAGCAGCGCCCGAGCCCC 3836
Db |||||||
QY 3837 TTTCAGGCCACAGCTGCCCGCAGACTCGCTGCAGCTCAGTCCCTCATCTGTAAGGTGAA 3896
Db |||||||
QY 701 TTCAGGCCACG--NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 644
```

```
QY 3897 GGGTATCGCAGATATGCTGACAGGAACAGTCTGTGGATGGAATGATGATCACTGCTAAGG 3956
Db |||||||
QY 643 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 584
Db |||||||
QY 3957 AAAGCAGCAGAGAGAGAGCG-TCCGGCGCCCGAGCCCACTATCACTGTGTCAGCGGTGCGG 4015
Db |||||||
QY 583 AAAGCAGCAGAGAGAGAGCGTCCGGCACCCAGAGCCCACTGTGACGCTCCAGCGTGGCGG 524
Db |||||||
QY 4016 TTCCCGCAGCAGCAGCTCAGCATCACTGACACTCACCTGACCTGCCCTGCCCTGGCGCAGAG 4075
Db |||||||
QY 523 CCCCCCAGCAGCAGC-----TCACACGAACACTCACCTGCCCTGCCCTGGCGGAG 470
Db |||||||
QY 4076 GGTACTCGCGAGCGGCACTTTGCACTCTGATGACCTCAAGACACTTTTCATGGCTGCCCTCT 4135
Db |||||||
QY 469 GGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 410
Db |||||||
QY 4136 GGCAGGGCAGCGCAGGCGAGTGACACTGTAGGAGCATAGCAAGCCAGGAGATGGGGTGA 4195
Db |||||||
QY 409 -----GGCAGGGCAGGTGAGTACTCTGTAGAGC--ANNNNNNNNNNNNNNNNNNNNNN 357
Db |||||||
QY 4196 GGGACACAGTCTTGAGCTGTCCACATGCATGTGCTCTCAAACTCTTCCAGATTTCTC 4255
Db |||||||
QY 356 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 297
Db |||||||
QY 4256 TAAGATAGCAGCCCTTCCCATTTGCCAGCTTAGCTCTTCCCGAGGGAGCTACT 4315
Db |||||||
QY 296 T---AATAGCAACACTTCCCATTTGCCAGCTTAGCTCTTCCCGAGGGAGCTACT 240
Db |||||||
QY 4316 CAGGACTCAGCTAGCATTAATCACTGTAATCGTCA-GGGGGTGTCTGTAGCTCAA 4374
Db |||||||
QY 239 CAGACTCATGTAGCATTAAGTCACTGTGAATTTGTAGGGGGGTGTGAGCTCA 180
Db |||||||
QY 4375 CCTCTGGGCGAGGGGAGCGCAGACTCCGTGGGAGAGCTCATTTCCACATTTGCCAA 4434
Db |||||||
QY 179 CTTCTGGGCGGGGAGCGCGAGACTCTTTGGGAGAGCTCATCCCCACCTTTGCCAA 120
Db |||||||
QY 4435 GACAGCTTTGTCAGCTGTCCACATTTGAGTCAGCTCTCCCGGGGAGAGCCCCGGC 4494
Db |||||||
QY 119 GACAGCTTTGTCAGCTGTCCACAGGAGCCAGACTGCTCCCGGGGAGAGAGCCCCGGC 60
Db |||||||
QY 4495 CCCCAGCACATAAAGAACTGCAGGCTTGGTACTGCAGAGTCTGGGTTGTAGAGAACTCTT 4554
Db |||||||
QY 59 CCCCAGCACATAAAGAACTGCAGGCTTGGTACTGC--AGTGTGGTTGTAGAGAACTCTT 2

RESULT 14
US-60-680-544-11056
; Sequence 11056, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE REFERENCE: 21590290000
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 11056
; LENGTH: 761
; TYPE: DNA
; ORGANISM: Macaca Fascicularis
; NAME/KEY: misc_feature
; LOCATION: (1)...(761)
; OTHER INFORMATION: n = A,T,C or G
```





Job time : 1775.13 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:00 ; Search time 144.36 Seconds  
(without alignments)  
2537.148 Million cell updates/sec

Title: US-09-155-676B-7

Perfect score: 5052

Sequence: 1 MAVNMACPGAGSGVGGQK.....PDGSFAWSVRKHGQLENRP 947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	5052	100.0	947	2	AAW42402 Amino aci
2	5052	100.0	947	5	AAU80371 Human cel
3	5052	100.0	947	7	ADP45075 Human kin
4	5052	100.0	947	8	ADS88253 Human pro
5	5045	99.9	947	2	AAI31665
6	5044	99.8	947	2	AAW82497 Human NIK
7	5044	99.8	947	2	AAW81564
8	5044	99.8	947	2	AAW81561
9	5040	99.8	947	7	AAI31667
10	5040	99.8	947	7	ABM85491 Human pro
11	5032	99.6	947	2	AAI31668
12	3439.5	68.1	697	8	ADP29813
13	2481.5	49.1	548	7	ABM85490 Mouse pro
14	2301	45.5	477	8	ADQ67717 Novel hum
15	1703	33.7	324	7	ABR61616 Human NIK
16	1703	33.7	324	7	ABR61608 Human NIK
17	1101	21.8	213	2	AAI31666
18	479	9.5	94	3	AAAG02579 Human sec
19	446	8.8	81	7	ABR61615
20	446	8.8	81	7	ABR61607 Human NIK
21	367	7.3	250	5	ABB78798 Serine/th
22	367	7.3	256	5	ADI17263 Polypepti
23	367	7.3	256	5	ADI17346 Polypepti
24	367	7.3	256	5	ADI17320 Polypepti
25	367	7.3	256	5	ADI17340 Polypepti

## ALIGNMENTS

## RESULT 1

AAW42402  
ID AAW42402 standard; protein; 947 AA.

AC AAW42402;

XX 15-APR-1998 (first entry)  
DT

XX Amino acid sequence of NF-kappaB inducing kinase.  
DE

XX Human tumour necrosis factor receptor-associated factor 2; TRAF2;  
KW TRAF-2 binding protein; NF-kappaB activity; NF-kappaB induction;

KW intracellular signalling activity; acute hepatitis;  
KW autoimmune-induced cell death.

XX Homo sapiens.

XX W09737016-AI.

PN 09-OCT-1997.

XX 01-APR-1997; 97WO-IL000117.

XX 02-APR-1996; 96IL-00117800.

PR 26-AUG-1996; 96IL-00119133.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX Wallach D, Malinin N, Boldin M, Kovalenko A, Mett I;

PI WPI; 1997-503101/46.

XX N-PSDB; AAV03325.

DR DNA encoding tumour necrosis factor receptor-associated factor binding

XX molecule - used for modulation or mediation in cells of the activity of

PT NF-kB.

XX Disclosure; Fig 6; 127pp; English.

XX The present sequence represents a NF-kappa inducing kinase (NIK). The

CC full length DNA sequence was obtained by PCR using clone 10 (AAV03324).

CC NIK was found to induce NF-kappaB even more effectively than TRAF2.

CC Proteins capable of binding to TRAF2 and NIK were identified. The NIK or

CC TRAF-2 binding proteins can be used for modulation or mediation in cells

CC of NF-kappaB activity or any other intracellular signalling activity

CC modulated or mediated by TRAF2. TRAF-binding proteins are especially used

CC for prevention or treatment of pathological conditions associated with NF

CC -kappaB induction, e.g. acute hepatitis, autoimmune-induced cell death,

26	360.5	7.1	467	4	AAE10314	AAE10314 Rat Tpl2
27	360	7.1	467	4	AAE10313	AAE10313 Human Tpl
28	360	7.1	467	7	ADJ94808	ADJ94808 Novel NOV
29	360	7.1	467	7	ADJ94796	ADJ94796 Novel NOV
30	360	7.1	468	7	ADJ94780	ADJ94780 Novel NOV
31	360	7.1	471	7	ADJ94790	ADJ94790 Novel NOV
32	360	7.1	474	7	ADJ94776	ADJ94776 Novel NOV
33	360	7.1	474	7	ADJ94798	ADJ94798 Novel NOV
34	360	7.1	475	7	ADJ94804	ADJ94804 Novel NOV
35	360	7.1	475	7	ADJ94788	ADJ94788 Novel NOV
36	360	7.1	478	7	ADJ94782	ADJ94782 Novel NOV
37	360	7.1	485	7	ADJ94794	ADJ94794 Novel NOV
38	360	7.1	486	7	ADJ94792	ADJ94792 Novel NOV
39	359.5	7.1	1493	8	ADQ94553	ADQ94553 Mouse ful
40	359	7.1	467	3	AAJ79244	AAJ79244 Human COT
41	359	7.1	467	4	AAE05951	AAE05951 Human cot
42	359	7.1	467	6	ADA05868	ADA05868 Human NOV
43	359	7.1	467	7	ADJ94774	ADJ94774 Novel NOV
44	359	7.1	467	8	ADI19737	ADI19737 Human tum
45	359	7.1	467	8	ADL82937	ADL82937 Human PRO

CC e.g. death of the beta Langerhans cells or the pancreas that results in  
CC diabetes, the death of cells in graft rejection, the death of  
CC oligodendrocytes in the brain in multiple sclerosis, and AIDS-inhibited T  
CC cell suicide which causes proliferation of the AIDS virus and hence the  
CC AIDS disease. The proteins are also useful for screening of ligands  
CC capable of binding to a protein, which are useful for modulating cellular  
CC activity modulated/mediated by TRAF2  
XX  
SQ Sequence 947 AA;

Query Match 100.0%; Score 5052; DB 2; Length 947;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60  
DB 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60  
QY 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSPFSEIFIAQESQYQSSESQIQINN 120  
DB 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSPFSEIFIAQESQYQSSESQIQINN 120  
QY 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQESCTIPVOE 180  
DB 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQESCTIPVOE 180  
QY 181 DESPLGAPVYRNTPOFTKPKELGQLCFKQIGELRPAALPRSELHKLISPLQCLNHVW 240  
DB 181 DESPLGAPVYRNTPOFTKPKELGQLCFKQIGELRPAALPRSELHKLISPLQCLNHVW 240  
QY 241 KLHHPODGGPLPLTHFFPYSLPHPPFPHPQPKPPLSFLGKLACVDSQKPLPDH 300  
DB 241 KLHHPODGGPLPLTHFFPYSLPHPPFPHPQPKPPLSFLGKLACVDSQKPLPDH 300  
QY 301 LSKLACVDSQKPLPGHLEPSCLSRGAHEKFSVEYLVHALQGSVSSQAHSLTSLAKTW 360  
DB 301 LSKLACVDSQKPLPGHLEPSCLSRGAHEKFSVEYLVHALQGSVSSQAHSLTSLAKTW 360  
QY 361 AARGSRSPSKPTEDNEGVLTAKLKPVDYREVEVHWATHQLRGSGFGEVHRMEDK 420  
DB 361 AARGSRSPSKPTEDNEGVLTAKLKPVDYREVEVHWATHQLRGSGFGEVHRMEDK 420  
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIRVPLYGAVREGPWNIWFELLEGSLGQL 480  
DB 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIRVPLYGAVREGPWNIWFELLEGSLGQL 480  
QY 481 VKEQCLPEDRALYILGQALEGLYHRSRIILHGVNDKADNVLSSDGSAAALCDFGHAVC 540  
DB 481 VKEQCLPEDRALYILGQALEGLYHRSRIILHGVNDKADNVLSSDGSAAALCDFGHAVC 540  
QY 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600  
DB 541 LQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600  
QY 601 RGFLCLKIASEPPPPVREIPPPSCAPLTAQAIQELRKEPIHRVSAALGKVNRLAQVGG 660  
DB 601 RGFLCLKIASEPPPPVREIPPPSCAPLTAQAIQELRKEPIHRVSAALGKVNRLAQVGG 660  
QY 661 LKSPRGVEYKEPRHPPPNQANVHQTALQAPRELSPRAPGPRPAEETTGAPKLQPLPPE 720  
DB 661 LKSPRGVEYKEPRHPPPNQANVHQTALQAPRELSPRAPGPRPAEETTGAPKLQPLPPE 720  
QY 721 PPEPNKSPPLTLKSESGMWEPLLSLEPAPARNPSPERKATVPPEOLOLELEFLN 780  
DB 721 PPEPNKSPPLTLKSESGMWEPLLSLEPAPARNPSPERKATVPPEOLOLELEFLN 780  
QY 781 SLSPFSLSEEQIILSCLSIDSLSDSEKNPFSKASQSSRDTLSSGVHSSSQAEARSS 840  
DB 781 SLSPFSLSEEQIILSCLSIDSLSDSEKNPFSKASQSSRDTLSSGVHSSSQAEARSS 840  
QY 841 SNMVLARGRPDTDSYNGVKVQIQLNGEHLHIREFHRVKVGDGIATGSIQIPAAAFS 900  
DB 841 SNMVLARGRPDTDSYNGVKVQIQLNGEHLHIREFHRVKVGDGIATGSIQIPAAAFS 900

QY 901 LVTKDQGPVRYDMEVPSDGSIDLOCTLAPDGSFANRWVKHGLENRPP 947  
DB 901 LVTKDQGPVRYDMEVPSDGSIDLOCTLAPDGSFANRWVKHGLENRPP 947  
RESULT 2  
ID AAU80371 standard; protein; 947 AA.  
XX AC AAU80371;  
XX DT 30-JUL-2002 (first entry)  
XX DE Human cellular kinase NIK protein.  
XX KW Human; virucide; cytomegalovirus infection; CMV; cellular kinase; RICK;  
XX RIP; Nck-Interacting kinase; NIK; MKK3; SRPK-2.  
XX OS Homo sapiens.  
XX PN EP1201765-A2.  
XX PD 02-MAY-2002.  
XX PF 15-OCT-2001; 2001EP-00124604.  
XX PR 16-OCT-2000; 2000US-0240750P.  
XX PA (AXXI-) AXKIMA PHARM AG.  
XX Schubart D, Habenberger P, Stein-Gerlach M, Bevec D;  
XX WPI; 2002-373930/41.  
XX DR N-PSDB; ABK51171.  
XX PT Identifying agents for treatment or prevention of cytomegalovirus  
XX infection, comprises contacting test compound with cellular kinase and  
XX detecting change in cellular kinase activity.  
XX PS Disclosure; Page 34-36; 49pp; English.  
XX The present invention relates to a new method for identifying compounds  
XX for treating and/or preventing cytomegalovirus (CMV) infection and/or  
XX related diseases. The method of the invention comprises contacting a test  
XX compound with at least one of the cellular kinases RICK, RIP, Nck-  
XX interacting kinase, MKK3 and SRPK-2 and detecting any change in kinase  
XX activity. The method of the invention can be used to treat and/or prevent  
XX CMV infections and related diseases. Oligonucleotides that can detect the  
XX specified kinases can also be used for diagnosis of infection. The  
XX present amino acid sequence represents the human cellular kinase NIK (Nck  
XX -interacting kinase) protein of the invention, as described above  
SQ Sequence 947 AA;  
Query Match 100.0%; Score 5052; DB 5; Length 947;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60  
DB 1 MAVMEMACPGAGSAGVQKELPKPKETPLGKQSSVYKLEAVEKSPVFCGKWEILND 60  
QY 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSPFSEIFIAQESQYQSSESQIQINN 120  
DB 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSPFSEIFIAQESQYQSSESQIQINN 120  
QY 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQESCTIPVOE 180  
DB 121 VAHATEGMARVCWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQESCTIPVOE 180  
QY 181 DESPLGAPVYRNTPOFTKPKELGQLCFKQIGELRPAALPRSELHKLISPLQCLNHVW 240

Db 181 DESPLGAPYVNTQFTKPLKEBGLGQLCFKQLEGLRPAIPRSELHKLISPLQCLNHVW 240  
Qy 241 KLHPDQGGPLPLPHTPPYSRLEPHFPFHPQLQPKPHPLESFLGKLACVDSQKPLPDPH 300  
Db 241 KLHPDQGGPLPLPHTPPYSRLEPHFPFHPQLQPKPHPLESFLGKLACVDSQKPLPDPH 300  
Qy 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360  
Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360  
Qy 361 AAGRSRSPKPTEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420  
Db 361 AAGRSRSPKPTEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420  
Qy 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIPMELLEGSLGQL 480  
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIPMELLEGSLGQL 480  
Qy 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVVKADNVLLSSDGSAAALCDFGHAVC 540  
Db 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVVKADNVLLSSDGSAAALCDFGHAVC 540  
Qy 541 LQPDGLGKSLITGDIIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLITGDIIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Qy 601 RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKYNRALQOVGG 660  
Db 601 RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKYNRALQOVGG 660  
Qy 661 LKSPWRGEYKEPRHPNPQANYHOTLHAQPRELSPRAPGPAPBETTGRAKPLQPLPPE 720  
Db 661 LKSPWRGEYKEPRHPNPQANYHOTLHAQPRELSPRAPGPAPBETTGRAKPLQPLPPE 720  
Qy 721 PPEPNKSPPLTLKSEESGMWELPLSLLEPAPARNPSPERKATVPQELQOLEIEFLN 780  
Db 721 PPEPNKSPPLTLKSEESGMWELPLSLLEPAPARNPSPERKATVPQELQOLEIEFLN 780  
Qy 781 SLSQPFSLLEEQEIILCSLSDSDSEKNPSKASOSSRDITLSSGVHSSQAARSS 840  
Db 781 SLSQPFSLLEEQEIILCSLSDSDSEKNPSKASOSSRDITLSSGVHSSQAARSS 840  
Qy 841 SNNVLARGRTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900  
Db 841 SNNVLARGRTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGSSQIPAAAFS 900  
Qy 901 LVTKDGPVRYDMEVPDSDIGDLOCTLAPDGSFANSWRVKGQLENRP 947  
Db 901 LVTKDGPVRYDMEVPDSDIGDLOCTLAPDGSFANSWRVKGQLENRP 947

## RESULT 3

ADP45075  
ID ADP45075 standard; protein; 947 AA.  
XX AC ADP45075;  
XX AC  
DT 12-FEB-2004 (first entry)  
XX Human kinase NIK.  
DE Human; protein kinase; enzyme; inhibitor; NIK.  
XX Homo sapiens.  
XX OS  
XX WO2003081210-A2.  
XX FN  
XX PD 02-OCT-2003.  
XX PF 20-MAR-2003; 2003WO-US008725.  
XX PR 21-MAR-2002; 2002US-0366892P.  
XX XX

PA (SUNE-) SUNESIS PHARM INC.

PI Prescott JC, Braisted A;

DR WPI; 2003-865136/80.

XX Identifying ligand binding to inactive conformation of target protein  
kinase (T) comprises contacting the conformation modified (T) which  
contains reactive group at binding site, with ligands and detecting  
kinase-ligand conjugate formation.

PS Disclosure; SEQ ID NO 44; 260pp; English.

XX The present invention relates to a method for identifying a ligand (L),  
which binds to an inactive conformation of target protein kinase (T). The  
method involves contacting inactive conformation of (T), which contains  
or is modified to contain a reactive group at or near a binding site of  
interest, with one or more ligand candidates capable of covalently  
bonding to the reactive group thus forming a kinase-(L) conjugate (C).  
The method is useful for identifying protein kinase inhibitors that  
preferentially bind to inactive conformation of a target protein kinase.  
The present sequence is a protein kinase which may be modified via an  
amino acid substitution, for use in the method of the invention.

SQ Sequence 947 AA;

Query Match 100.0%; Score 5052; DB 7; Length 947;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAVMEMACPGAPGAVGQKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGKWEILND 60

Db 1 MAVMEMACPGAPGAVGQKELPKPKETPLPKKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKESAGPAAISIIIAQAECSQSEFSTFSEIRFIAGSKQVSQESLQIINN 120

Db 61 VITKGTAKESAGPAAISIIIAQAECSQSEFSTFSEIRFIAGSKQVSQESLQIINN 120

Qy 121 VAHATEGMARVCKGKRRSKARKKRRKSSKSLAHAGVALAKPLPRTPQESCTIPVQE 180

Db 121 VAHATEGMARVCKGKRRSKARKKRRKSSKSLAHAGVALAKPLPRTPQESCTIPVQE 180

Qy 181 DESPLGAPYVNTQFTKPLKEBGLGQLCFKQLEGLRPAIPRSELHKLISPLQCLNHVW 240

Db 181 DESPLGAPYVNTQFTKPLKEBGLGQLCFKQLEGLRPAIPRSELHKLISPLQCLNHVW 240

Qy 241 KLHPDQGGPLPLPHTPPYSRLEPHFPFHPQLQPKPHPLESFLGKLACVDSQKPLPDPH 300

Db 241 KLHPDQGGPLPLPHTPPYSRLEPHFPFHPQLQPKPHPLESFLGKLACVDSQKPLPDPH 300

Qy 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360

Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAHEKFSVEYLVLHVALQGSVSSQAHSLSLAKTW 360

Qy 361 AAGRSRSPKPTEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420

Db 361 AAGRSRSPKPTEDNEGVLITLTKLPVDYVEEVEVHWATHQLRLGRSGFGEVHRMEDK 420

Qy 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIPMELLEGSLGQL 480

Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVLYGAVREGPWNIIPMELLEGSLGQL 480

Qy 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVVKADNVLLSSDGSAAALCDFGHAVC 540

Db 481 VKEQGCLPEDRALYVLGQALEGLYLSRRILHGDVVKADNVLLSSDGSAAALCDFGHAVC 540

Qy 541 LQPDGLGKSLITGDIIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600

Db 541 LQPDGLGKSLITGDIIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600

Qy 601 RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKYNRALQOVGG 660

Db 601 RGPLCLKIASPPPPVREIPSPCAPLTAQAIQEGRLKEPIHRVSAELGGKYNRALQOVGG 660





diagnostic; therapy.  
Homo sapiens.  
US5854003-A.  
29-DEC-1998.  
26-FEB-1998; 98US-00032475.  
03-JUL-1997; 97US-00887518.  
(TULA-) TULARIK INC.  
Wu L, Rothe M;  
WPI; 1999-094902/08.  
N-PSDB; AAV73917.  
Screening agents for modulating interaction of nuclear factor kappaB inducing kinase - with kinase-binding target, useful for controlling levels of the kinase, for treatment and diagnosis of conditions associated with e.g. inhibition of signal transduction by tumour necrosis factor.  
Claim 6; Col 13-20; 16pp; English.  
This sequence represents a novel human nuclear factor-kappaB (NFkB)-inducing kinase (NIK) which is used in a method for screening for agents that modulate the interaction of NIK with a NIK-binding target. This protein can be used as a modulator of cellular functions at the NIK level, or for development of such compounds. NFkB is involved in expression of many immune and inflammatory responses and of some important viral genes. The protein may be used diagnostically and therapeutically, in conditions associated with abnormal utilisation of pathways that involve NFkB, e.g. inhibition of signal transduction by tumour necrosis factor (TNF)

Query Match  
Best Local Similarity 99.8%; Score 5044; DB 2; Length 947;  
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAVMEMACGAPGSAVGQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACGAPGSAVGQKELPKAKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
QY 61 VITTKCTAKEGSEAGAAISIIAQAECENSQBFSTFSEIRIAGSKQYSQESLDQIPNN 120  
Db 61 VITTKCTAKEGSEAGAAISIIAQAECENSQBFSTFSEIRIAGSKQYSQESLDQIPNN 120  
QY 121 VAHATEGKMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPTPEQESCTIPVOE 180  
Db 121 VAHATEGKMARVCKWKGKRSKARKKRSKSLAHAGVALAKPLPTPEQESCTIPVOE 180  
QY 181 DESPLGAPYVNTPOFTKPLKEPGLQCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVNTPOFTKPLKEPGLQCFKQLEGRLPALPRSELHKLISPLQCLNHVW 240  
QY 241 KLHHPQDGGPLPLPHTPPYSRLPHPPFPPLQPKWHPLESFLGKLACVDSQKPLDPDH 300  
Db 241 KLHHPQDGGPLPLPHTPPYSRLPHPPFPPLQPKWHPLESFLGKLACVDSQKPLDPDH 300  
QY 301 LSKLACVDSQKPLPGLPCLSRGAHEKFSVEEYLVHALQGSVSSQASHLSLAKTW 360  
Db 301 LSKLACVDSQKPLPGLPCLSRGAHEKFSVEEYLVHALQGSVSSQASHLSLAKTW 360  
QY 361 AARGSRSEPSKTEDNEGVLTLTKLPVDYVEEVEVHWATHQLRGSGFGEVHRMEDK 420  
Db 361 AARGSRSEPSKTEDNEGVLTLTKLPVDYVEEVEVHWATHQLRGSGFGEVHRMEDK 420  
QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVLYGAVREGPWNIIFMELLEGGSLGQL 480

Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPIVLYGAVREGPWNIIFMELLEGGSLGQL 480  
QY 481 VKEQGCLPEDRALYYLGOALEGLEYLHHSRRILHGDVKADNVLLSSDGGSHAALCDFGHAVC 540  
Db 481 VKEQGCLPEDRALYYLGOALEGLEYLHHSRRILHGDVKADNVLLSSDGGSHAALCDFGHAVC 540  
QY 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVVLRGSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLLTGDYIPGTETHMAPEVVVLRGSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600  
QY 601 RGPCLCKIASPPPPVREIPPCAPLTAQAIQEGLRKEPIHVSAAELGKGNRALQQVGG 660  
Db 601 RGPCLCKIASPPPPVREIPPCAPLTAQAIQEGLRKEPIHVSAAELGKGNRALQQVGG 660  
QY 661 LKSPWRGEYKEPRHPPPPNQANYHOTLHAQPRELSPRAFPRAETTTGRAPKLQPLPPE 720  
Db 661 LKSPWRGEYKEPRHPPPPNQANYHOTLHAQPRELSPRAFPRAETTTGRAPKLQPLPPE 720  
QY 721 PPEPNKSPPLTLSKEESGMWEPFLSSLEPAPARNPSSPERKATVPPEOLOOLBIEIPLN 780  
Db 721 PPEPNKSPPLTLSKEESGMWEPFLSSLEPAPARNPSSPERKATVPPEOLOOLBIEIPLN 780  
QY 781 SLSQPFSLSEOEQILSCLSIDSLSDSEKNPKSKASQSSRDTLSGVSWSQAEARSS 840  
Db 781 SLSQPFSLSEOEQILSCLSIDSLSDSEKNPKSKASQSSRDTLSGVSWSQAEARSS 840  
QY 841 SNNMVLARGRPTDTPSYFNGVKVQISLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Db 841 SNNMVLARGRPTDTPSYFNGVKVQISLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
QY 901 LVTXKDGQPVRYDMEVPSGIDLOCTIAPDGSFAWSWRVKGQLENRP 947  
Db 901 LVTXKDGQPVRYDMEVPSGIDLOCTIAPDGSFAWSWRVKGQLENRP 947

RESULT 7  
AAW81564  
ID AAW81564 standard; protein; 947 AA.  
AC AAW81564;  
XX AAW81564;  
DT 04-FEB-1999 (first entry)  
XX Human NF-kB-inducing kinase (NIK) polypeptide.  
DE NIK; IkB; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;  
KW tumour necrosis factor; binding; genetic hybridisation; screening;  
KW signal transduction; biopharmaceutical; human.  
XX Homo sapiens.  
XX US5843721-A.  
XX 01-DEC-1998.  
XX 03-JUL-1997; 97US-00887518.  
XX 03-JUL-1997; 97US-00887518.  
XX (TULA-) TULARIK INC.  
XX Wu L, Rothe M;  
XX WPI; 1999-044580/04.  
XX N-PSDB; AAV71603.  
XX Probe, vector or recombinant nucleic acid encoding a polypeptide,  
PT especially human nuclear factor kappa-B-inducing kinase protein - useful  
PT for producing recombinant protein.  
XX Claim 7; Col 13-18; 15pp; English.  
XX



CC This represents a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)  
CC polypeptide. The NIK polypeptide or its fragment has one or more  
CC activities selected from kinase activity and inhibitory activity; Ikb  
CC kinase-alpha and beta binding activity and inhibitory activity;  
CC tumour necrosis factor (TNF) receptor-associated factor 2 binding  
CC activity and binding inhibitory activity; Ikb binding activity and  
CC binding inhibitory activity; NF-kB activating and inhibitory activity. A  
CC vector containing the NIK nucleic acid can be used to transform host  
CC cells for the recombinant production of the protein. The NIK nucleic acid  
CC and the polypeptide may be used in diagnosis (e.g. genetic hybridisation  
CC screen for NIK transcripts), therapy (e.g. NIK kinase inhibitors to  
CC inhibit TNF signal transduction), and in the biopharmaceutical industry  
XX  
SQ Sequence 947 AA;  
Query Match 99.8%; Score 5044; DB 2; Length 947;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MAVMEMACPGAGSAVGQKQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMEMACPGAGSAVGQKQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQSFSTFSERIFIAGSKOYSQSESIDQIPNN 120  
Db 61 VITKGTAKGSEAGPAAISIIAQACENSQSFSTFSERIFIAGSKOYSQSESIDQIPNN 120  
Qy 121 VAHATEGKMARVCWKGRKRSKARKKRSKSLAHAGVALAKPLRTPQESCTIIPVOE 180  
Db 121 VAHATEGKMARVCWKGRKRSKARKKRSKSLAHAGVALAKPLRTPQESCTIIPVOE 180  
Qy 181 DESPLGAPYVRNTPQFTKPLKEPGLGQCFKQKQEGRLPALPRSELHKLISPLQCLNHVW 240  
Db 181 DESPLGAPYVRNTPQFTKPLKEPGLGQCFKQKQEGRLPALPRSELHKLISPLQCLNHVW 240  
Qy 241 KLHPHQDGGPLPLTPHPFYSRLPHPPFHPPLQPKWKHPHLESFLGKLCVDSQKPLDPDH 300  
Db 241 KLHPHQDGGPLPLTPHPFYSRLPHPPFHPPLQPKWKHPHLESFLGKLCVDSQKPLDPDH 300  
Qy 301 LSKLACVDSPKPLPGPHLEPSCLSRGAEKFSVEEYLVAHALQGVSSQASHLTSIAKTW 360  
Db 301 LSKLACVDSPKPLPGPHLEPSCLSRGAEKFSVEEYLVAHALQGVSSQASHLTSIAKTW 360  
Qy 361 AARGSRREPSPKTNEGVLLTEKLPVDYEEVEVHWATHQLRGRGSGFGEVHRMEDK 420  
Db 361 AARGSRREPSPKTNEGVLLTEKLPVDYEEVEVHWATHQLRGRGSGFGEVHRMEDX 420  
Qy 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWVNTFMELLEGSSLGQL 480  
Db 421 QTGFQCAVKVRLEVFRAEELMACAGLTSPRIVPLYGAVREGPWVNTFMELLEGSSLGQL 480  
Qy 481 VKEOGCLPEORALYVLGOALEGLVYLSRRITLHGDVKNVLLSSDGSAAALCDFGHAVC 540  
Db 481 VKEOGCLPEORALYVLGOALEGLVYLSRRITLHGDVKNVLLSSDGSAAALCDFGHAVC 540  
Qy 541 LQPDGLGKSLITGDIYIPGTETHMAPEVVLGRSCDAKVDVMSCCMHLMLHNGCHPWTQFF 600  
Db 541 LQPDGLGKSLITGDIYIPGTETHMAPEVVLGRSCDAKVDVMSCCMHLMLHNGCHPWTQFF 600  
Qy 601 RGPLCLKIASPPVRIIPSCAPLTAAQIQEGLRKSPPIHRVSAAEILGGKVNRAALQVGG 660  
Db 601 RGPLCLKIASPPVRIIPSCAPLTAAQIQEGLRKSPPIHRVSAAEILGGKVNRAALQVGG 660  
Qy 720 LKSPWRGEYKEPRPPPNQANYHOTLHAQRELSPRAGPRPAEETTGRAKLPQPLPPE 720  
Db 720 LKSPWRGEYKEPRPPPNQANYHOTLHAQRELSPRAGPRPAEETTGRAKLPQPLPPE 720  
Qy 721 PPRENKSPPLTLKSEESGMWEPPLSSLEPAPARNPSSPERKATVPQELQQLIEFLN 780  
Db 721 PPRENKSPPLTLKSEESGMWEPPLSSLEPAPARNPSSPERKATVPQELQQLIEFLN 780  
Qy 781 SLSQFFSLEEQEQLSCLSDLSLSDSEKNPSKASQSRDRTLSSGVHWSWSSQAEARSS 840  
XX

Db 781 SLSQFFSLEEQEQLSCLSDLSLSDSEKNPSKASQSRDRTLSSGVHWSWSSQAEARSS 840  
Qy 841 SNNVVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
XX  
Db 841 SNNVVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAATGISSQIPAAAFS 900  
Qy 901 LVTKGQGPVRYDMEVPSGIDLOCTTLAPDGSFAWSRVKIGQLENRP 947  
XX  
Db 901 LVTKGQGPVRYDMEVPSGIDLOCTTLAPDGSFAWSRVKIGQLENRP 947  
RESULT 8  
AAW81561  
ID AAW81561 standard; protein; 947 AA.  
XX  
AC AAW81561;  
XX  
DT 04-FEB-1999 (first entry)  
XX  
Human NF-kB-inducing kinase (NIK) polypeptide.  
XX  
NIK; Ikb; NF-kB; TNF; kinase; nuclear factor kappa B; inhibition;  
XX tumour necrosis factor; binding; genetic hybridisation; screening;  
XX signal transduction; biopharmaceutical; immunogen; pharmacological;  
XX transcription regulator; human.  
XX  
OS Homo sapiens.  
XX  
XX US5844073-A.  
XX  
PD 01-DEC-1998.  
XX  
PF 13-FEB-1998; 98US-00023321.  
XX  
PR 03-JUL-1997; 97US-00887518.  
XX  
XX (TULA-) TULARIK INC.  
XX  
PI Wu L, Rothe M;  
XX  
XX WPI; 1999-044664/04.  
DR N-PSDB; AAV69285.  
XX  
New isolated peptide comprising a specified 947 amino acid sequence - has  
PT e.g. kinase activity, kinase inhibitory activity, Ikb kinase-alpha  
PT binding activity, and Ikb kinase-alpha binding inhibitory activity.  
XX  
XX Claim 1; Col 13-18; 15pp; English.  
XX  
PS This represents a nuclear factor kappa B (NF-kB)-inducing kinase (NIK)  
CC polypeptide. The NIK polypeptide or its fragment has one or more  
CC activities selected from kinase activity and inhibitory activity; Ikb  
CC kinase-alpha and beta binding activity and binding inhibitory activity;  
CC tumour necrosis factor (TNF) receptor-associated factor 2 binding  
CC activity and binding inhibitory activity; Ikb binding activity and  
CC binding inhibitory activity; NF-kB activating and inhibitory activity.  
CC The NIK nucleic acid and the polypeptide may be used in diagnosis (e.g.  
CC genetic hybridisation screen for NIK transcripts), therapy (e.g. NIK  
CC kinase inhibitors to inhibit TNF signal transduction), and in the  
CC biopharmaceutical industry (e.g. as immunogens, reagents for isolating  
CC other transcription regulators, and reagents for screening chemical  
CC libraries for pharmacological agents)  
XX  
SQ Sequence 947 AA;  
Query Match 99.8%; Score 5044; DB 2; Length 947;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 946; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MAVMEMACPGAGSAVGQKQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60  
XX  
Db 1 MAVMEMACPGAGSAVGQKQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60

```
QY 61 VITKGTAKGSEAGAPAAISIIAQACENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
Db 61 VITKGTAKGSEAGAPAAISIIAQACENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
QY 121 VAHATEGMARVCMWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVQE 180
Db 121 VAHATEGMARVCMWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVQE 180
QY 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQGLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQGLGEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPQDGGPLPLTHPPFPYVSRPLPFPPLPQPKPPLPESFLGKLACVDSQKPLPDH 300
Db 241 KLHHPQDGGPLPLTHPPFPYVSRPLPFPPLPQPKPPLPESFLGKLACVDSQKPLPDH 300
QY 301 LSKLACVDSQKPLPQGHLEPSCLSRGNAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
Db 301 LSKLACVDSQKPLPQGHLEPSCLSRGNAHEKFSVEEYLVHALQGSVSSQAHSLSLAKTW 360
QY 361 AARGSRSPKTEDNEGVLTEKLPVDYREVEVHWATHQLRGRSGFGEVHRMEDX 420
Db 361 AARGSRSPKTEDNEGVLTEKLPVDYREVEVHWATHQLRGRSGFGEVHRMEDX 420
QY 421 QTCFQCAVKKVLEVFRAELMACAGLTSPRIPLYGAVREGPVNFI FMELEGGSLGQL 480
Db 421 QTCFQCAVKKVLEVFRAELMACAGLTSPRIPLYGAVREGPVNFI FMELEGGSLGQL 480
QY 481 VKEQGLCPEDRALYILGQALEGLEVYHSRIILHGDVKNLSSDGSAAALCDFGHAVC 540
Db 481 VKEQGLCPEDRALYILGQALEGLEVYHSRIILHGDVKNLSSDGSAAALCDFGHAVC 540
QY 541 LQPDGLGKSLLDYIPGTETHTMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
Db 541 LQPDGLGKSLLDYIPGTETHTMAPEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600
QY 601 RGPLCLKIASPPVREIPSPCAPTAQAIQGLRKEPIHRVSAELGCKVNRALQQVGG 660
Db 601 RGPLCLKIASPPVREIPSPCAPTAQAIQGLRKEPIHRVSAELGCKVNRALQQVGG 660
QY 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPCPRAEPTTGRAPKLPPLPPE 720
Db 661 LKSPMRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPCPRAEPTTGRAPKLPPLPPE 720
QY 721 PEPNKSPLTLKSKEGWEPLPLSSLEPAPARNPSPERKATVPEQELQQLRIELFLN 780
Db 721 PEPNKSPLTLKSKEGWEPLPLSSLEPAPARNPSPERKATVPEQELQQLRIELFLN 780
QY 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKASOSSRDTLSSGVHSSWSSQAEARSS 840
Db 781 SLSQPFSLSEEQILSCLSIDSLSDSEKNPKASOSSRDTLSSGVHSSWSSQAEARSS 840
QY 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGTSIQIPAAAFS 900
Db 841 SNMVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIAITGTSIQIPAAAFS 900
QY 901 LVTKDQGVVRDMEVDSGIDLQCTLAPDGSFAWSWRVKGQLENRP 947
Db 901 LVTKDQGVVRDMEVDSGIDLQCTLAPDGSFAWSWRVKGQLENRP 947

RESULT 9
AAV31667
ID AAV31667 standard; protein; 947 AA.
XX
XX
AC AAV31667;
XX
XX
DT 09-NOV-1999 (first entry)
XX
DE NIK kinase-deficient mutant (T559A).
XX
KW NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;
interleukin-1; cytokine; inhibitor; anti-inflammatory; apoptosis;
KW
```

```
KW anti-apoptotic; human; mutant.
XX
OS Homo sapiens.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 559 /note= "wild-type Thr is substituted by Ala"
FT
XX
XX WO9943704-Al.
XX
XX 02-SEP-1999.
XX
XX 25-FEB-1999; 99WO-US004110.
XX
XX 27-FEB-1999; 98US-0076299P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Greene WC, Lin X, Gelezuinas R;
XX WPI; 1999-518837/43.
XX
XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent
immune response, e.g. an inflammatory response or an anti-apoptotic
response.
XX
XX Claim 8; Page; 48pp; English.
XX
XX The present sequence represents a kinase-deficient mutant of human NF-kB
inducing kinase (NIK), in which the threonine residue at position 559 of
the wild-type protein (see AAY31665) is replaced by alanine. The
invention relates to: (a) an N-terminus deletion mutant NIK protein (see
AAY31666); and (b) a kinase deficient NIK mutant protein, especially the
present sequence of the sequence given in AAY31668, that inhibits auto-
phosphorylation or transphosphorylation. The invention provides the
molecular basis for cytokine induction of NF-kB-dependent immune and
inflammatory responses, emphasising a role for both NIK-NIK and NIK-IKK
(IkB-specific kinase) interactions. A novel and highly specific method
for modulating NF-kB-dependent immune, inflammatory and anti-apoptotic
responses is based on interruption of the critical protein interaction of
NIK and IKK. The mutant NIK proteins are used in claimed methods for
inhibiting NF-kB-dependent gene expression. The kinase-deficient NIK
mutant proteins inhibit the activation and phosphorylation of endogenous
IKK. Note: The present sequence is not shown in the specification but is
derived from the wild-type NIK sequence given in figure 10
XX
XX Sequence 947 AA;
```

```
Query Match 99.8%; Score 5040; DB 2; Length 947;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 945; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAYWEMACGAPGASVQKQKELPKPKETPPGLKQKQSVYKLEAVEKSPVFCGKWEILND 60
Db 1 MAYWEMACGAPGASVQKQKELPKPKETPPGLKQKQSVYKLEAVEKSPVFCGKWEILND 60
QY 61 VITKGTAKGSEAGAPAAISIIAQACENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
Db 61 VITKGTAKGSEAGAPAAISIIAQACENSQEFSTFSEIFIAAGSKQYSQESLDOI PNN 120
QY 121 VAHATEGMARVCMWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVQE 180
Db 121 VAHATEGMARVCMWKGKRSKARKKRSKSLAHAGVALAKPLRTPPEQSCCTIPVQE 180
QY 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQGLGEGRLPALPRSELHKLISPLQCLNHVW 240
Db 181 DESPLGAPYVRNTPOFTKPLKEPGLGQLCFKQGLGEGRLPALPRSELHKLISPLQCLNHVW 240
QY 241 KLHHPQDGGPLPLTHPPFPYVSRPLPFPPLPQPKPPLPESFLGKLACVDSQKPLPDH 300
Db 241 KLHHPQDGGPLPLTHPPFPYVSRPLPFPPLPQPKPPLPESFLGKLACVDSQKPLPDH 300
```



QY 721 PPEPNKSPPLTLSESGMWEPLSLSELPAPARNPSSPERKATVPEQELQOIEIFLNLN 780  
 Db 721 PPEPNKSPPLTLSESGMWEPLSLSELPAPARNPSSPERKATVPEQELQOIEIFLNLN 780  
 QY 781 SLSPFSLERQEQILSLCLSDSLSDSEKKNPFSKASQSSRDTLSSGVHSSWSQAARSS 840  
 Db 781 SLSPFSLERQEQILSLCLSDSLSDSEKKNPFSKASQSSRDTLSSGVHSSWSQAARSS 840  
 QY 841 SNWVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDITATGISSQIPAAAFS 900  
 Db 841 SNWVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDITATGISSQIPAAAFS 900  
 QY 901 LVTKDQOPVRYDMEVDSGIDLOCTLAPDGSFAWSWRVXHGQLENRP 947  
 Db 901 LVTKDQOPVRYDMEVDSGIDLOCTLAPDGSFAWSWRVXHGQLENRP 947

## RESULT 11

AAV31668  
 ID AAY31668 standard; protein; 947 AA.

AC AAY31668;

DT 09-NOV-1999 (first entry)

XX NIK kinase-deficient mutant (S549A,T552A,T559A).

KW NIK; NF-kB inducing kinase; inflammation; tumour necrosis factor;  
 KW interleukin-1; cytokine; inhibitor; antiinflammatory; apoptosis;  
 KW anti-apoptotic; human; mutant.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 549 /note= "wild-type Ser is substituted by Ala"

FT Misc-difference 552 /note= "wild-type Thr is substituted by Ala"

FT Misc-difference 559 /note= "wild-type Thr is substituted by Ala"

XX WO943704-A1.

XX 02-SEP-1999.

XX 25-FEB-1999; 99WO-US004110.

XX 27-FEB-1998; 98US-0076299P.

XX (REGC ) UNIV CALIFORNIA.

XX Greene WC, Lin X, Gelezuinas R;

XX WPI; 1999-518837/43.

XX New NF-kB inducing kinase proteins, used for inhibiting a kB-dependent  
 PT immune response, e.g. an inflammatory response or an anti-apoptotic  
 PT response.

XX Claim 9; Page; 48pp; English.

XX The present sequence represents a kinase-deficient mutant of human NF-kB  
 CC inducing kinase (NIK), in which the serine residue at position 549 of the  
 CC wild-type protein (see AAY31665) is replaced by alanine, the threonine  
 CC residue at position 552 is replaced by alanine, and the threonine residue  
 CC at position 559 is replaced by alanine. The invention relates to: (a) an  
 CC N-terminus deletion mutant NIK protein. (see AAY31666); and (b) a kinase  
 CC deficient NIK mutant protein, especially the present sequence of the  
 CC sequence given in AAY31667, that inhibits auto-phosphorylation or  
 CC transphosphorylation. The invention provides the molecular basis for  
 CC cytokine induction of NF-kB-dependent immune and inflammatory responses,

CC emphasising a role for both NIK-NIK and NIK-IKK (IkB-specific kinase)  
 CC interactions. A novel and highly specific method for modulating NF-kB-  
 CC dependent immune, inflammatory and anti-apoptotic responses is based on  
 CC interruption of the critical protein interaction of NIK and IKK. The  
 CC mutant NIK proteins are used in claimed methods for inhibiting NF-kB-  
 CC dependent gene expression. The kinase-deficient NIK mutant proteins  
 CC inhibit the activation and phosphorylation of endogenous IKK. Note: The  
 CC present sequence is not shown in the specification but is derived from  
 CC the wild-type NIK sequence given in figure 10

XX Sequence 947 AA;

Query Match 99.6%; Score 5032; DB 2; Length 947;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 943; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAVMEMACPGAGPGSAVGQOKELPKPKETPLPKKKQSSVYKLEAVEKSPVFCGKWEILND 60

Db 1 MAVMEMACPGAGPGSAVGQOKELPKPKETPLPKKKQSSVYKLEAVEKSPVFCGKWEILND 60

QY 61 VITKTAKEGSEAGPAAIISIIAQACENSOQSFSTPFSERIFIAGSKOYSQSESIDQIPNN 120

Db 61 VITKTAKEGSEAGPAAIISIIAQACENSOQSFSTPFSERIFIAGSKOYSQSESIDQIPNN 120

QY 121 VAHATEGMARVCWKGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCITIPVQE 180

Db 121 VAHATEGMARVCWKGRSKARKKKKSKSLAHAGVALAKPLPRTPEQESCITIPVQE 180

QY 181 DESPLGAPYVRNTPQFTKPLKEGQLCPKQLEGRLPALPRSELHKLSPLOCLNHVW 240

Db 181 DESPLGAPYVRNTPQFTKPLKEGQLCPKQLEGRLPALPRSELHKLSPLOCLNHVW 240

QY 241 KLHPDQGGPLPLTPHPFYSLRPLPFPPLPQPKPHPLESFLGKLACVDSQKPLDPH 300

Db 241 KLHPDQGGPLPLTPHPFYSLRPLPFPPLPQPKPHPLESFLGKLACVDSQKPLDPH 300

QY 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQASHLTSIAKTW 360

Db 301 LSKLACVDSQKPLPGPHLEPSCLSRGAHEKFSVEEYLVHALQGSVSSQASHLTSIAKTW 360

QY 361 AAGRSRSPSKPTEDNEGVLLETKLPVDYREEVHMAHQLRLGRGSGFGEVHRMEDK 420

Db 361 AAGRSRSPSKPTEDNEGVLLETKLPVDYREEVHMAHQLRLGRGSGFGEVHRMEDK 420

QY 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIVPLYGAVREGFWNIFMELLEGGSLGQL 480

Db 421 QTGFQCAVKVRLEVFRAELMACAGLTSPRIVPLYGAVREGFWNIFMELLEGGSLGQL 480

QY 481 VKEQGCLPEDRALYILGQALEGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540

Db 481 VKEQGCLPEDRALYILGQALEGLEYLHSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540

QY 541 LQPDGLGKSLITGDYIPGTETHMADEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600

Db 541 LQPDGLGKSLITGDYIPGTETHMADEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF 600

QY 601 RGPLCLKLIASEPPPVREIPSCAPLTAQAOIEGLRKEPIHRVSAAEELGKVNALQOVGG 660

Db 601 RGPLCLKLIASEPPPVREIPSCAPLTAQAOIEGLRKEPIHRVSAAEELGKVNALQOVGG 660

QY 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGPRPAEETTGAPKLQPLPPE 720

Db 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGPRPAEETTGAPKLQPLPPE 720

QY 721 PPEPNKSPPLTLSESGMWEPLSLSELPAPARNPSSPERKATVPEQELQOIEIFLNLN 780

Db 721 PPEPNKSPPLTLSESGMWEPLSLSELPAPARNPSSPERKATVPEQELQOIEIFLNLN 780

QY 781 SLSPFSLERQEQILSLCLSDSLSDSEKKNPFSKASQSSRDTLSSGVHSSWSQAARSS 840

Db 781 SLSPFSLERQEQILSLCLSDSLSDSEKKNPFSKASQSSRDTLSSGVHSSWSQAARSS 840

QY 841 SNWVLARGRPTDTPSYFNGVKVQIQSLNGEHLHIREFHRVKVGDITATGISSQIPAAAFS 900

Db 841 SNNWVLARGRTDPSYNGVKVQLQSLNGEHLHIREFHRVVKVGIATGSSQIPAAAFS 900  
Qy 901 LVTKDGPVRYDMEVPDSDGLDQCTLAPDGSFAMSRVKGQLENRP 947  
Db 901 LVTKDGPVRYDMEVPDSDGLDQCTLAPDGSFAMSRVKGQLENRP 947

RESULT 12  
ADP29813  
ID ADP29813 standard; protein; 697 AA.

XX AC ADP29813;

XX DT 12-AUG-2004 (first entry)

XX DE Human secreted protein SEQ ID #580.

XX KW Cytostatic; Antinflammatory; Immunosuppressive; Antibacterial; Virucide;  
XX KW cancer; Inflammatory; Immune; human secreted protein.

XX OS Homo sapiens.

XX PN WO2004035732-A2.

XX PD 29-APR-2004.

XX PF 28-AUG-2003; 2003WO-US026780.

XX PR 29-AUG-2002; 2002US-0406576P.

XX PR 29-AUG-2002; 2002US-0406579P.

XX PR 29-AUG-2002; 2002US-0406585P.

XX PR 29-AUG-2002; 2002US-0406588P.

XX PR 29-AUG-2002; 2002US-0406608P.

XX PR 29-AUG-2002; 2002US-0406611P.

XX PR 29-AUG-2002; 2002US-0406612P.

XX PR 29-AUG-2002; 2002US-0406616P.

XX PR 29-AUG-2002; 2002US-0406640P.

XX PR 29-AUG-2002; 2002US-0406642P.

XX PR 29-AUG-2002; 2002US-0406646P.

XX PR 29-AUG-2002; 2002US-0406653P.

XX PR 29-AUG-2002; 2002US-0406655P.

XX PR 29-AUG-2002; 2002US-0406666P.

XX PR 17-SEP-2002; 2002US-0410946P.

XX PR 17-SEP-2002; 2002US-0410947P.

XX PR 17-SEP-2002; 2002US-0410948P.

XX PR 17-SEP-2002; 2002US-0410949P.

XX PR 17-SEP-2002; 2002US-0410953P.

XX PR 17-SEP-2002; 2002US-0410957P.

XX PR 17-SEP-2002; 2002US-0410958P.

XX PR 17-SEP-2002; 2002US-0410959P.

XX PR 17-SEP-2002; 2002US-0410960P.

XX PR 17-SEP-2002; 2002US-0410961P.

XX PR 17-SEP-2002; 2002US-0410962P.

XX PR 17-SEP-2002; 2002US-0411019P.

XX PR 17-SEP-2002; 2002US-0411022P.

XX PR 17-SEP-2002; 2002US-0411023P.

XX PR 17-SEP-2002; 2002US-0411024P.

XX PR 17-SEP-2002; 2002US-0411032P.

XX PR 17-SEP-2002; 2002US-0411035P.

XX PR 17-SEP-2002; 2002US-0411037P.

XX PR 17-SEP-2002; 2002US-0411041P.

XX PR 17-SEP-2002; 2002US-0411045P.

XX PR 17-SEP-2002; 2002US-0411046P.

XX PR 17-SEP-2002; 2002US-0411048P.

XX PR 17-SEP-2002; 2002US-0411052P.

XX PR 17-SEP-2002; 2002US-0411055P.

XX PR 17-SEP-2002; 2002US-0411073P.

XX PR 17-SEP-2002; 2002US-0411082P.

XX PR 17-SEP-2002; 2002US-041101P.

XX PR 17-SEP-2002; 2002US-0411111P.

XX PR 18-APR-2003; 2003US-0463700P.

XX PR 18-APR-2003; 2003US-0463708P.

PR 18-APR-2003; 2003US-0463716P.  
PR 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467201P.  
PR 02-MAY-2003; 2003US-0467203P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 22-MAY-2003; 2003US-0472420P.  
PR 22-MAY-2003; 2003US-0472430P.  
PR 09-JUN-2003; 2003US-0476609P.  
PR 09-JUN-2003; 2003US-0476641P.  
PR 08-JUL-2003; 2003US-0485218P.  
PR 08-JUL-2003; 2003US-0485223P.  
PR 08-JUL-2003; 2003US-0485224P.  
PR 08-JUL-2003; 2003US-0485325P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 15-JUL-2003; 2003US-0486891P.  
PR 15-JUL-2003; 2003US-0486960P.  
PR 08-AUG-2003; 2003US-0493341P.  
PR 08-AUG-2003; 2003US-0493370P.  
PR 08-AUG-2003; 2003US-0493573P.  
PR 08-AUG-2003; 2003US-0493577P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
XX  
XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;  
PI Halenbeck RP, Huang WM, Kothakota S, Haishan L, Linnemann T;  
PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;  
XX  
XX WPI; 2004-348438/32.

XX New nucleic acid molecule for diagnosing, preventing or treating diseases  
XX such as proliferative (e.g. cancer), inflammatory, immune, metabolic,  
XX genetic, bacterial and viral diseases.

XX Claim 1; SEQ ID NO 1811; 428pp; English.

XX The present invention relates to an isolated nucleic acid molecule  
XX encoding a polypeptide which is believed to be cytostatic,  
XX antiinflammatory, immunosuppressive, antibacterial and virucidal. The  
XX composition and methods are useful for diagnosing, preventing and  
XX treating diseases such as proliferative (e.g. cancer), inflammatory,  
XX immune, metabolic, genetic, bacterial and viral diseases. The present  
XX sequence represents a human secreted protein. The present sequence is  
XX available on WIPOWEB and is not in the specification.

XX SQ Sequence 697 AA;

Query Match 68.1%; Score 3439.5; DB 8; Length 697;  
Best Local Similarity 72.7%; Pred. No. 5.8e-246;  
Matches 688; Conservative 0; Mismatches 2; Indels 257; Gaps 5;

Qy 1 MAVNEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60

Db 1 MAVNEMACPGAGSAGVQKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGKWEILND 60

Qy 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSTFSERIFIAGSKQYSQESLDQIPNN 120

Db 61 VITKGTAKGSEAGPAAISIIAQACENSQEFSTFSERIFIAGSKQYSQESLDQIPNN 120

Qy 121 VAHATEGMARVCKWKKRRSKARKKSKSLAHAGVALAKPLPTPEGECTIPVQ 180

Db 121 VAHATEGMARVCKWKKRRSKARKKSKSLAHAGVALAKPLPTPEGECTIPVQ- 179

Qy 181 DESPLGAPYVRNTPQFTKPLKELGQLCFKQGLGRLPALPRSELHKLISPLQCLNHW 240

Db 180 ----- 179

Qy 241 KLHHPDQGGPLPLTPHPFPYSRLPHPPFHPQWKPHPLESFLGKLCACVDSQKPLDPH 300

Db 180 ----- 179

QY 301 LSKLACVDSKPLPGPHLEPCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW 360  
 Db 180 -----LPGPHLEPCLSRGAHEKFSVEEYLVHALQGSVSSQAHSLTSLAKTW 227  
 QY 361 AARGSRSPSPKTEDEGVLTEKLKPDYDEYREEVHWATHQLRLGRSGFGEVHRMEDK 420  
 Db 228 AARGSRSPSPKTEDEGVLTEKLKPDYDEYREEVHWATHQLRLGRSGFGEVHRMEDK 287  
 QY 421 QTGFQCAVKVRLVFRABELMACAGLTSPRIPLVYCAVREGPWNI FMELEGGSLGOL 480  
 Db 288 QTGFQCAVKK-----YCAVREGPWNI FMELEGGSLGOL 322  
 QY 481 VKEQGLCPEDRALYILGOALEGLEYLHSRRLHGDVKADNVLSSDGHAAALCDFGHAVC 540  
 Db 323 VKEQGLCPEDRALYILGOALEGLEYLHSRRLHGDVKADNVLSSDGHAAALCDFGHAVC 382  
 QY 541 LPDGLGKSLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600  
 Db 383 LPDGLGKSLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 442  
 QY 601 RGPLCLKIASPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAABELGCKVNRALQOVGG 660  
 Db 443 RGPLCLKIASPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAABELGCKVNRALQOVGG 499  
 QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGRPAEETTCRAPKLQPLPPE 720  
 Db 500 -----E 500  
 QY 721 PPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSPERKATVPEQLOOLELFLN 780  
 Db 501 PPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSPERKATVPEQLOOLELFLN 557  
 QY 781 SLQSPFSLQEQILSCLSDSLSLSDSEKNPASKASQSSRDTLSSGVHSSQAEARSS 840  
 Db 558 SLQSPFSLQEQILSCLSDSLSLSDSEKNPASKASQSSRDTLSSGVHSSQAEARSS 588  
 QY 841 SNNMVLARGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVKGVDIATGISSQIPAAAFS 900  
 Db 589 -----LARGRPDTSYFNGVKVQIQSLNGEHLHIREFHRVKGVDIATGISSQIPAAAFS 643  
 QY 901 LVTKDGPVRYDMEVPDGSIDLOCTLAPDGSFAMSWRVKHGQLENRP 947  
 Db 644 LVTKDGPVRYDMEVPDGSIDLOCTLAPDGSFAMSWRVKHGQLENRP 690

RESULT 13  
 ABM85490  
 ID ABM85490 standard; protein; 548 AA.  
 AC ABM85490;  
 XX 18-NOV-2004 (first entry)  
 DT Mouse protein sequence mCP14606.  
 DE Cytostatic; carcinoma; lymphoma; cancer; murine.  
 XX Mus musculus.  
 OS Mus musculus.  
 FN WO2003073826-A2.  
 XX 12-SEP-2003.  
 XX 28-FEB-2003; 2003WO-US006235.  
 XX 01-MAR-2002; 2002US-00087192.  
 PA (SAGR-) SAGRES DISCOVERY.  
 PI Morris DW;  
 DR WPI; 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
 PT comprises a nucleotide sequence.  
 XX Claim 5; SEQ ID NO 885; Opp; English.  
 PS The present invention relates to novel DNA and protein sequences which  
 XX are associated with carcinomas. The sequences are useful for: (i) for  
 CC screening drug candidates; (ii) for screening of bioactive agent capable  
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
 CC carcinoma including lymphoma. The present sequence is one such CAP. Note:  
 CC This patent is an equivalent to basic patent US2002182586A1, for which no  
 CC sequence data was published  
 XX Sequence 548 AA;  
 SQ

Query Match 49.1%; Score 2481.5; DB 7; Length 548;  
 Best Local Similarity 86.1%; Pred. No. 4.9e-175;  
 Matches 466; Conservative 32; Mismatches 36; Indels 7; Gaps 2;  
 QY 407 GRGSFGEVHRMEDKOTGFQCAVKVRLVFRABELMACAGLTSPRIPLVYCAVREGPWNI 466  
 Db 15 QSSSFGFVHRMKDQTFQCAVKVRLVFRABELMACAGLTSPRIPLVYCAVREGPWNI 74  
 QY 467 IFMELLEGGSLGOLVKEQGLCPEDRALYILGOALEGLEYLHSRRLHGDVKADNVLSSD 526  
 Db 75 IFMELLEGGSLGOLVKEQGLCPEDRALYILGOALEGLEYLHSRRLHGDVKADNVLSSD 134  
 QY 527 GSHAALCDFGHAVCLOPDGLGKSLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMW 586  
 Db 135 GSRAALCDFGHAVCLOPDGLGKSLTGDIYIPGTETHMAPEVLGRSCDAKVDVWSSCCMW 194  
 QY 587 LHMLNGCHPWTQFPRGPKLKIASEPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAAE 646  
 Db 195 LHMLNGCHPWTQFPRGPKLKIASEPPVREIPSPCAPLTAQAIQGLRKEPIHRVSAAE 254  
 QY 647 LGKVNALQOVGKLSKPEGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGRPAEET 706  
 Db 255 LRRKVGKALQOVGKLSKPEGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGRPAEET 307  
 QY 707 TGRAPKLQPLPPEPPPEPNKSPPLTLKSKEGSMWEPLPLSSLEPAPARNPSPERKATVP 766  
 Db 308 TDGAPEQPPLPPEPPPEPNKSPALNLSKESGTWEPLPLSSLDPAKAGSFDDRATLP 367  
 QY 767 EQELOQLEIEFLNLSLQSPFSLQEQILSCLSDSLSLSDSEKNPASKASQSSRDTLSS 826  
 Db 368 EQELOQLEIEFLNLSLQSPFSLQEQILSCLSDSLSLSDSEKNPASKASQSSRDTLSS 427  
 QY 827 GVHSSWSQAEARSSNNMVLARGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVKGDI 886  
 Db 428 GVHSSWSQAEARSSNNMVLARGRPTDTSYFNGVKVQIQSLNGEHLHIREFHRVKGDI 487  
 QY 887 ATGISSQIPAAAFSLVTKDGPVRYDMEVPDGSIDLOCTLAPDGSFAMSWRVKHGQLENR 946  
 Db 488 ATGISSQIPATFSLVTKDGPVRYDMEVPDGSIDLOCTLAPDGSFAMSWRVKHGQLENR 547  
 QY 947 P 947  
 Db 548 P 548  
 RESULT 14  
 ADQ67717  
 ID ADQ67717 standard; protein; 477 AA.  
 XX ADQ67717;  
 AC

XX 07-OCT-2004 (first entry)  
XX Novel human protein sequence #2383.  
XX osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;  
XX gene therapy; diagnostic marker; morbid state; osteoporosis;  
XX neurological disease; Alzheimer's disease; Parkinson's disease; dementia;  
XX cancer.  
XX Homo sapiens.  
XX EPI440981-A2.  
XX 28-JUL-2004.  
XX 21-JAN-2004; 2004EP-00001196.  
XX 21-JAN-2003; 2003JP-00102206.  
XX 09-MAY-2003; 2003JP-00131392.  
XX (REAS-) RES ASSOC BIOTECHNOLOGY.  
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
XX Yamamoto J, Isono Y, Nagai K, Irie R;  
XX WPI; 2004-535376/52.  
XX N-PSDB; ADQ67410.  
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,  
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.  
XX Claim 1; SEQ ID NO 4878; 2449pp; English.  
XX The invention relates to 2495 novel polynucleotides (I) and their encoded  
XX polypeptides, sequences hybridizing to these nucleotides, sequences  
XX encoding partial polypeptides and sequences having 70% or 90% identity to  
XX the nucleotide and protein sequences. The nucleotides and polypeptides  
XX are useful as diagnostic markers or therapeutic target for the diseases  
XX or morbid states. They are also useful for treating osteoporosis,  
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,  
XX dementia and various cancers. This sequence corresponds to a protein  
XX sequence of the invention.  
XX Sequence 477 AA;  
Query Match 45.5%; Score 2301; DB 8; Length 477;  
Best Local Similarity 99.3%; Pred. No. 9.8e-162;  
Matches 431; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 514 GDVRADNVLLSSDGSAAALCDFGHAVCLQPPGLGKSLTGTGYPCTETHMAPEVVLGRSC 573  
DB 44 GGLAANDNVLLSSDGSAAALCDFGHAVCLQPPGLGKSLTGTGYPCTETHMAPEVVLGRSC 103  
QY 574 DAKVDVSSCCMLHMLNGCHPWTQFRGPICLKIASEPPPPVRIIPSPCAPLTAQIQEG 633  
DB 104 DAKVDVSSCCMLHMLNGCHPWTQFRGPICLKIASEPPPPVRIIPSPCAPLTAQIQEG 163  
QY 634 LRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYHQTLLHAQREL 693  
DB 164 LRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYHQTLLHAQREL 223  
QY 694 SPRAPGRPAETTCRAPKLPPLPPPEPKNPPLTLKESGMEPLPLSLLEPAPA 753  
DB 224 SPRAPGRPAETTCRAPKLPPLPPPEPKNPPLTLKESGMEPLPLSLLEPAPA 283  
QY 754 RNPSSPERKATVPQEQLOQLIEFLNLSLSPFSLSEQQLSLCLSDLSLSDSEKNP 813  
DB 284 RNPSSPERKATVPQEQLOQLIEFLNLSLSPFSLSEQQLSLCLSDLSLSDSEKNP 343  
QY 814 SKAQSSRDTLSSGVHSSWSSQAEARSSSNMVLARGRPTDTPSYFNGVKVQIQSLNGEHL 873  
DB 344 SKAQSSRDTLSSGVHSSWSSQAEARSSSNMVLARGRPTDTPSYFNGVKVQIQSLNGEHL 403

QY 874 HIREFHRVKVGDIAITGISSQIPAAAFSLVTKDGPVRYDMEVPDSDGLQCTLAPDGSFA 933  
DB 404 HIREFHRVKVGDIAITGISSQIPAAAFSLVTKDGPVRYDMEVPDSDGLQCTLAPDGSFA 463  
QY 934 WSWRVKVGQLENRP 947  
DB 464 WSWRVKVGQLENRP 477  
RESULT 15  
ABR61616  
ID ABR61616 standard; protein; 324 AA.  
XX AC ABR61616;  
XX 15-JAN-2004 (first entry)  
XX Human NIK C-terminal fragment (residues 624-947).  
XX NIK; IL-2; interleukin 2; common gamma chain; cgammac; antiinflammatory;  
XX gastrointestinal; antiarthritic; antirheumatic; osteopathic; cardiant;  
XX antiasthmatic; nootropic; neuroprotective; antiarteriosclerotic; NIKK;  
XX immunosuppressive; antithyroid; nuclear factor KB inducing kinase;  
XX gene therapy; antianemic; NIK kinase; human.  
XX Homo sapiens.  
XX WO2003087380-A1.  
XX 23-OCT-2003.  
XX 15-APR-2003; 2003WO-IL000317.  
XX 18-APR-2002; 2002IL-00149217.  
XX 08-OCT-2002; 2002IL-00152183.  
XX (YEDA ) YEDA RES & DEV CO LTD.  
XX Wallach D, Ramakrishnan P, Shmushkovich T;  
XX WPI; 2003-845333/78.  
XX New nuclear factor inducing kinase or its mutein, variant, fusion  
XX protein, functional derivative, circularly permuted derivative or  
XX fragment, useful for treating an autoimmune disease, infarct, Alzheimer's  
XX disease or atherosclerosis.  
XX Claim 2; Page 92-93; Opp; English.  
XX The invention relates to a NIK (nuclear factor (NF)-kB-inducing kinase)  
XX or its mutein, variant, fusion protein, functional derivative, circularly  
XX permuted derivative or fragment. Specific antibodies and small  
XX molecules capable of modulating the interaction between interleukin 2 (IL  
XX -2) common gamma chain (cgammac) and NIK kinase (NIKK) are useful for the  
XX manufacture of a medicament for the treatment of a disease, e.g. a  
XX disease resulting from excessive immune response such as rheumatoid  
XX arthritis, osteoarthritis, inflammatory bowel disease, asthma, cardiac  
XX infarct, Alzheimer's disease or atherosclerosis; or an autoimmune disease  
XX such as immune thyroiditis, or other arthropathies, such as autoimmune  
XX haemolytic anemia. The small molecule is useful for modulating signaling  
XX through cgammac. The present sequence represents a human NIK fragment  
XX Sequence 324 AA;  
Query Match 33.7%; Score 1703; DB 7; Length 324;  
Best Local Similarity 100.0%; Pred. No. 1.3e-117;  
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 624 PLTAAQIQEGLRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYH 683  
DB 1 PLTAAQIQEGLRKEPIHRVSAELGGKVNALQOVGGLKSPWRGEYKEPRHPPPNQANYH 60

QY 684 QTLHAQPRELSPRACPRPAEATTGRAPKLQPPLPPEPPENKSPPLTLKSKESGMWEP 743  
Db |||||  
61 QTLHAQPRELSPRACPRPAEATTGRAPKLQPPLPPEPPENKSPPLTLKSKESGMWEP 120  
QY 744 PLSSLEPAPARNPSPERKATVPEQELQOLEIELEFNLNSLQPFSLBEOQIILSCLSIDSL 803  
Db |||||  
121 PLSSLEPAPARNPSPERKATVPEQELQOLEIELEFNLNSLQPFSLBEOQIILSCLSIDSL 180  
QY 804 SLSDDEKNPKSKASOSSRDTLSSGVHWSQAEARSSWNMYLARGRPTDTPSYFNGVKV 863  
Db |||||  
181 SLSDDEKNPKSKASOSSRDTLSSGVHWSQAEARSSWNMYLARGRPTDTPSYFNGVKV 240  
QY 864 QIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFSLVTKDQPVRYDMEVPDGSIDIQ 923  
Db |||||  
241 QIQSLNGEHLHIREFHRVKVGDIAIGISSQIPAAAFSLVTKDQPVRYDMEVPDGSIDIQ 300  
QY 924 CTLAPDGSFAMSWRVKHGQLENRP 947  
Db |||||  
301 CTLAPDGSFAMSWRVKHGQLENRP 324

Search completed: June 7, 2005, 11:57:28  
Job time : 150.36 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:40:40 ; Search time 138.585 Seconds  
(without alignments)  
3499.208 Million cell updates/sec

Title: US-09-155-676B-7

Perfect score: 5052

Sequence: 1 MAVNEMACPGAGSAGVQK.....PDGSFAWSRVKQGLENRP 947

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5052	100.0	947	1 M3KE_HUMAN	Q99558 homo sapien
2	5040	99.8	947	2 Q8IYN1	Q8IYN1 homo sapien
3	4196.5	83.1	942	1 M3KE_MOUSE	Q9WU16 mus musculu
4	2301	45.5	477	2 Q6ZM21	Q6ZM21 homo sapien
5	2110	41.8	433	2 Q68D39	Q68D39 homo sapien
6	360	7.1	467	1 M3K8_HUMAN	P41279 homo sapien
7	360	7.1	467	2 Q6FG25	Q6FG25 homo sapien
8	357.5	7.1	467	1 M3K8_MOUSE	Q07174 mus musculu
9	356.5	7.1	467	1 M3K8_RAT	Q63562 rattus norv
10	356	7.0	1291	2 Q9WT82	Q9WT82 mus musculu
11	355.5	7.0	1493	1 M3K1_MOUSE	P33349 mus musculu
12	353	7.0	619	1 M3K2_MOUSE	Q61083 mus musculu
13	353	7.0	1493	1 M3K1_RAT	Q62925 rattus norv
14	351	6.9	1158	2 Q6C2G1	Q6C2G1 varrowia li
15	350.5	6.9	619	2 Q9NYK3	Q9NYK3 homo sapien
16	346	6.8	423	2 Q8NC32	Q8NC32 homo sapien
17	346	6.8	1495	1 M3K1_HUMAN	Q13233 homo sapien
18	345	6.8	618	1 M3K2_HUMAN	Q9Y2U5 homo sapien
19	343.5	6.8	614	2 Q6RFY3	Q6RFY3 lycopersico
20	341.5	6.8	944	2 Q6C1T1	Q6C1T1 varrowia li
21	341.5	6.8	944	2 Q6Z8J5	Q6Z8J5 varrowia li
22	341	6.7	1320	2 Q9UVC9	Q9UVC9 emericella
23	338.5	6.7	611	2 Q8RFY4	Q8RFY4 nicotiana b
24	335.5	6.6	690	2 Q8H334	Q8H334 oryza sativ
25	332.5	6.6	894	2 Q7XU11	Q7XU11 oryza sativ
26	332	6.6	735	2 Q6Z189	Q6Z189 oryza sativ
27	327	6.5	709	2 Q7XV58	Q7XV58 oryza sativ
28	326.5	6.5	651	1 M3K3_ARATH	Q2042 arabidopsis
29	325	6.4	651	1 M3K2_ARATH	Q9FZ36 arabidopsis
30	325	6.4	1288	2 Q75PK3	Q9FZ36 homo sapien
31	325	6.4	1478	1 BCK1_YEAST	Q01389 saccharomyc

#### RESULT 1

ID	M3KE_HUMAN	STANDARD;	PRT;	947 AA.
AC	Q99558;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	Mitogen-activated protein kinase kinase 14 (EC 2.7.1.37) (NP-kappa beta-inducing kinase) (Serine/threonine-protein kinase NIK) (HsNIK)			
DE	Name=MAP3K14; Synonyms=NIK;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF 429-LYS-LYS-430.			
RX	MEDLINE=97172277; PubMed=9020361;			
RA	Malinin N.L., Boldin M.P., Kovalenko A.V., Wallach D.;			
RT	"MAP3K-related kinase involved in NF-kappaB induction by TNF, CD95 and IL-1.";			
RL	Nature 385:540-544 (1997).			
RN	[2]			
RP	INTERACTION WITH PELI3.			
RX	MEDLINE=22756745; PubMed=12874243;			
RA	Jensen L.E., Whitehead A.S.;			
RT	"Pellino3, a novel member of the Pellino protein family, promotes activation of c-Jun and Elk-1 and may act as a scaffolding protein.";			
RL	J. Immunol. 171:1500-1506(2003).			
CC	-1- FUNCTION: Lymphotoxin beta-activated kinase which seems to be exclusively involved in the activation of NF-kappa-B and its transcriptional activity. Induces the processing of NF-kappa-B 2/P100. Could act in a receptor-selective manner (By similarity).			
CC	-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.			
CC	-1- SUBUNIT: Binds to TRAF2, TRAF5, TRAF6, IKKA and NF-kappa-B 2/P100 (By similarity). Interacts with PELI3.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-1- TISSUE SPECIFICITY: Weakly expressed in testis, small intestine, spleen, thymus, peripheral blood leukocytes, prostate, ovary and colon.			
CC	-1- PTM: Autophosphorylated.			
CC	-1- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP kinase kinase subfamily.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).			
CC	-----			
CC	EMBL; Y10256; CAA71306.1; -.			
DR	HSSP; P71584; 1MRU.			

DR IntAct; Q9558; --  
DR Genew; HGNC:6853; MAP3K14.  
DR MTM; 604655; --  
DR InterPro; IPR011009; Kinase\_like.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR Pfam; PF00069; Pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS0108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Phosphorylation; Serine/threonine-protein kinase;  
KW Transferase.  
FT DOMAIN 400 655 Protein kinase.  
FT NP\_BIND 406 414 ATP (By similarity).  
FT BINDING 429 429 ATP (By similarity).  
FT ACT\_SITE 515 515 Proton acceptor (By similarity).  
FT MUTAGEN 429 430 KK->AA: Loss of autophosphorylation.  
SQ SEQUENCE 947 AA; 104097 MW; D307F76993BC9661 CRC64;  
Query Match 100.0%; Score 5052; DB 1; Length 947;  
Best Local Similarity 100.0%; Pred. No. 6.5e-241;  
Matches 947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAVMENACPGAGSAGVQOKELPKPKETPLPKGKQSSVYKLEAVEKSPVFCGKWEILND 60  
Db 1 MAVMENACPGAGSAGVQOKELPKPKETPLPKGKQSSVYKLEAVEKSPVFCGKWEILND 60  
QY 61 VITKGTAKGSEAGPAAISIIIAQACENSQFSPFSEIFIAQKQYQSSESLQIPNN 120  
Db 61 VITKGTAKGSEAGPAAISIIIAQACENSQFSPFSEIFIAQKQYQSSESLQIPNN 120  
QY 121 VAHATGKWARVCKGKRSKARKKPKKSSKSLAHAGVALAKPLRTPQECSCTIPVQE 180  
Db 121 VAHATGKWARVCKGKRSKARKKPKKSSKSLAHAGVALAKPLRTPQECSCTIPVQE 180  
QY 181 DESPLGAPVVRNTPQTKLPKPGQLGKLFKQEGRLPRSELKHLISPLQCLNHVW 240  
Db 181 DESPLGAPVVRNTPQTKLPKPGQLGKLFKQEGRLPRSELKHLISPLQCLNHVW 240  
QY 241 KLHPDQGGPLPLTHPPFYSRLPHPPHPLQPKWPHPLESLGKLACVDSQKPLPDHP 300  
Db 241 KLHPDQGGPLPLTHPPFYSRLPHPPHPLQPKWPHPLESLGKLACVDSQKPLPDHP 300  
QY 301 LSKLACVDSQKPLPGHPLPSCLSGAHEKFSVEEVLVHALQGSVSSQASHLSLAKTW 360  
Db 301 LSKLACVDSQKPLPGHPLPSCLSGAHEKFSVEEVLVHALQGSVSSQASHLSLAKTW 360  
QY 361 AARGSRSPKPTEDNEGVLTEKLPVDYREYREVEVHWATHQLRLGRGSFGVEHRMEDK 420  
Db 361 AARGSRSPKPTEDNEGVLTEKLPVDYREYREVEVHWATHQLRLGRGSFGVEHRMEDK 420  
QY 421 QTGFQCAVKVRLVFRABELMACAGLTSPTRIPLVYGAVREGPWNIIFMELLEGSLQGL 480  
Db 421 QTGFQCAVKVRLVFRABELMACAGLTSPTRIPLVYGAVREGPWNIIFMELLEGSLQGL 480  
QY 481 VKEQGLCPEDRALYLQGLALEGLVYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540  
Db 481 VKEQGLCPEDRALYLQGLALEGLVYLSRRILHGDVKADNVLLSSDGSAAALCDFGHAVC 540  
QY 541 LQPDGLGKSLGADYIPGTETHMAFEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600  
Db 541 LQPDGLGKSLGADYIPGTETHMAFEVVLGRSCDAKVDVWSSCCMWLHMLNGCHPWTQFF 600  
QY 601 RGPCLCKTASEPPVRETPPSCAPITQAQIQGLRKEPIHRSVAELGKGNRALQQVGG 660  
Db 601 RGPCLCKTASEPPVRETPPSCAPITQAQIQGLRKEPIHRSVAELGKGNRALQQVGG 660  
QY 661 LKSPWRGEYKEPRHPNPNQANVHQLHAQPRELSPRAPGPRPAETTTGRPKLQPLPPE 720  
Db 661 LKSPWRGEYKEPRHPNPNQANVHQLHAQPRELSPRAPGPRPAETTTGRPKLQPLPPE 720  
QY 721 PPEPNKSPPLTLUSKESGMEPLPLSSLEPAPARNPSPERKATVPEQELQLELFLN 780

Db 721 PPEPNKSPPLTLUSKESGMEPLPLSSLEPAPARNPSPERKATVPEQELQLELFLN 780  
QY 781 SLSQPSFLEEQILSLCLSIDSLSDSEKKNPSKASOSSRDTLSSGVHWSQAEARSS 840  
Db 781 SLSQPSFLEEQILSLCLSIDSLSDSEKKNPSKASOSSRDTLSSGVHWSQAEARSS 840  
QY 841 SNMVLARGRPDTPSYFNGVKVQIQLNGEHLHIREFHRVKVGDIAIGSIQIPAAAFS 900  
Db 841 SNMVLARGRPDTPSYFNGVKVQIQLNGEHLHIREFHRVKVGDIAIGSIQIPAAAFS 900  
QY 901 LVTKDQGPVRYDMEVDSGIDLQCTLAPDGSFAMVRVKHGOLENRP 947  
Db 901 LVTKDQGPVRYDMEVDSGIDLQCTLAPDGSFAMVRVKHGOLENRP 947  
RESULT 2  
Q8IYN1 PRELIMINARY; PRT; 947 AA.  
AC Q8IYN1; OS Homo sapiens (Human);  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
DE MAP3K14 protein.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.U., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RC SEQUENCE FROM N.A.  
RC TISSUE=Lymph;  
RA Strausberg R.;  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC035576; AAHJ5576.1; -;  
DR HSP; F71584; 106Y.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004672; F:protein kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006488; F:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR011009; Kinase-like.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
DR Pfam; PF00069; Pkinase; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS0108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 947 AA; 104041 MW; C9D10F67FF7F48AC CRC64;

Query Match	99.8%; Score 5040; DB 2; Length 947;	
Best Local Similarity	99.8%; Pred. NO. 2.5e-240;	
Matches 945; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1 MAVMEMCPGAPGSAVGOKELPKPKETPLGKKQSSVYKLEAVEKSPVFCGWEILND 60	
Db	1 MAVMEMCPGAPGSAVGOKELPKAKETPLGKKQSSVYKLEAVEKSPVFCGWEILND 60	
Qy	61 VITKGTAKGSEAGPAAISIIAQAECENSQSFSTFSERIFIAQSKQYSESIDQIPNN 120	
Db	61 VITKGTAKGSEAGPAAISIIAQAECENSQSFSTFSERIFIAQSKQYSESIDQIPNN 120	
Qy	121 VAHATEGMARVCKWKKRRSKARKKRRKKSKSLAHAGVALAKPLPRTPEQESTIPVQE 180	
Db	121 VAHATEGMARVCKWKKRRSKARKKRRKKSKSLAHAGVALAKPLPRTPEQESTIPVQE 180	
Qy	181 DESPLGAPYVNTPOFTKPEGLGOLCFKQEGELRPALPRSELHKLISPLQCLNHVW 240	
Db	181 DESPLGAPYVNTPOFTKPEGLGOLCFKQEGELRPALPRSELHKLISPLQCLNHVW 240	
Qy	241 KLHPQDGGPLPPLTHPPYSLRPLPFPFPLQPKWHPLESFLGKLACVDQSKPLDPDPH 300	
Db	241 KLHPQDGGPLPPLTHPPYSLRPLPFPFPLQPKWHPLESFLGKLACVDQSKPLDPDPH 300	
Qy	301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVALQGVSSQASHLSLAKTW 360	
Db	301 LSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVALQGVSSQASHLSLAKTW 360	
Qy	361 AARGSRSEPKTNEGVLITTEKLKPDVDEYEEVHWATHQRLRGSGFGEVHRMEDK 420	
Db	361 AARGSRSEPKTNEGVLITTEKLKPDVDEYEEVHWATHQRLRGSGFGEVHRMEDK 420	
Qy	421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIVLYGAVREGPWNIIFMELLEGGSLGOL 480	
Db	421 QTGFQCAVKVRLEVFRAEELMACAGLTSPIVLYGAVREGPWNIIFMELLEGGSLGOL 480	
Qy	481 VKEQGLPEDRALYVLGQALEGLYLSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540	
Db	481 VKEQGLPEDRALYVLGQALEGLYLSRRILHGDVADNVLSSDGSAAALCDFGHAVC 540	
Qy	541 LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTQFF 600	
Db	541 LQPDGLKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMHLMLNGCHPWTQFF 600	
Qy	601 RGPLCLKTASPPPPVREIIPPCAPLTAQAIQEGLRKEPIHRVSAEELGGKVNRAALQOVGG 660	
Db	601 RGPLCLKTASPPPPVREIIPPCAPLTAQAIQEGLRKEPIHRVSAEELGGKVNRAALQOVGG 660	
Qy	661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPAPBETTTGRAPKLQPLPPE 720	
Db	661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPRAPGPAPBETTTGRAPKLQPLPPE 720	
Qy	721 PPEPNKSPPLTLSKEESGMEPLPLSSLEPAPARNPSSPERKATVPPEOQLQELIPLN 780	
Db	721 PPEPNKSPPLTLSKEESGMEPLPLSSLEPAPARNPSSPERKATVPPEOQLQELIPLN 780	
Qy	781 SLSQPFSLSEOEQILSLCLSIDLSLSDSEKNPKASQSSRDTLSSGVHSSQAEARSS 840	
Db	781 SLSQPFSLSEOEQILSLCLSIDLSLSDSEKNPKASQSSRDTLSSGVHSSQAEARSS 840	
Qy	841 SNNVLAGRTRDTPSYFNGVKVQIQLNGELHLIREFHRVKVGDIAITGSSQIPAAAFS 900	
Db	841 SNNVLAGRTRDTPSYFNGVKVQIQLNGELHLIREFHRVKVGDIAITGSSQIPAAAFS 900	
Qy	901 LVTXGQVRYDMEVPSGIDLQCTLAPDGSFANWRVKGQLENRP 947	
Db	901 LVTXGQVRYDMEVPSGIDLQCTLAPDGSFANWRVKGQLENRP 947	

RESULT 3

M3KE MOUSE

ID M3KE\_MOUSE

STANDARD; PRT; 942 AA.

Query Match

Best Local Similarity

Matches 797; Conservative

83.1%; Score 4196.5; DB 1; Length 942;

84.0%; Pred. No. 9.3e-199;

53; Mismatches 90; Indels 9; Gaps 4;

Qy	1	MAWMACPGAPGSAVQKQKELPKPKBKTPLGKKOSSVYKLEAVKSPVFCGKWEILND	60
Db	1	MAWMEVACPGTPGSAVQKQKELAKAKBKTQSLGKKQSCIFKLEAVKSPVFCGKWEILND	60
Qy	61	VITKGTAKGSEAGPAAISIIIAQECENSQEFSPFTFSEIFITAGSKQYQSQESLDOI	120
Db	61	VITKGTAKDSEGGPPAIIIAQECENSQEFSPFTFSEIFITAGSKQYQSQESLDOI	120
Qy	121	VAHATGKKWAVCWKGRSKARKKRRKKSLSLAHAGVALAKPLPRTPEQBSCTTPVQE	180
Db	121	VAHATGKKWAVCCRRKHGKARKKRRKKSLSLAQAGVALAKPLPRTPEQBSCTTPVQE	180
Qy	181	DESPGAPVVRNTPOPTKLPKPLGQCLCFKOLGELRPALESELHKLISPLQCLNHVW	240
Db	181	DESPGLNLYARNVSQFTKPLGFGPLGHLCFKKQDEGLRPVLPPELUHKUISPLQCLNHVW	240
Qy	241	KLHHQDGGPLPLPHTPPYSRLPHPPFPHPIQPKPHPLES-FLGKLACVDSQKPLP-D	298
Db	241	KLHHQATGPRPHPTPPFYSQMPHPPFPYPLEPWKPYMLDSAVLDKLAGVSGQRPLPGP	300
Qy	299	PHLSKLACVDSKPLPGPHLEPSCLSRGAHEKFSVEEYLVAHQSVSSQAHSTUSLAK	358
Db	301	PHLSQAHGDSQKPLPGPHLES CSPRGALEKVPVEEYLVAHQSVSSQAHSLASTAK	360
Qy	359	TWAAGRSRSPKTEDENEGVLLETKLPVDYREEVHMAATHOLRLGRGSFGEVHME	418
Db	361	TWSSGSAKLQGLPETEDNEGVLLETKLPVDYREEVHWHHTHPRVGRGSFGEVHMK	420
Qy	419	DKQTGFOCAVKVRLEVFRAEELMACAGITSPRIVPLYCAVREGPWNI FMELEGGSLG	478
Db	421	DKQTGFOCAVKVRLEVFREELVACAGLSSPRIVPLYCAVREGPWNI FMELEGGSLG	480
Qy	479	QLVKEQGLCPEDRALYILGOALEGEYLHRSRIILHGDVKANDVLISSGSHAALCDFGHA	538
Db	481	QLIKQMGCLPEDRALYILGOALEGEYLHTRILHGDVKANDVLISSDGSRAALCDFGHA	540
Qy	539	VCLQPDGLGKSLLTGDYIPGTETHMAPEVILGRSCDAKVDVWSSCCMMHLMNGCHPWQ	598
Db	541	LCLQPDGLGKSLLTGDYIPGTETHMAPEVWVGKPCDAKVDI WSSCCMMHLMNGCHPWQ	600
Qy	599	FFRGPLCLKIASEPPVREI PPSCAPLTAQAOEGLRKEPTHRVSAAELGKVNRALQOV	658
Db	601	YFRGPLCLKIASEPPPIREI PPSCAPLTAQAOEGLRKEPVHRSAMELRKVKQALQEV	660
Qy	659	GGLXSPMRGEYKEPRHPPPNQANYHTLHAQPRELSPRAPGPRAEETTGRAKLOPLP	718
Db	661	GGLXSPWKGEYKEPRPPQDQATCHQTLPTPRE-NP-----PAKANTDGAPEPQPLP	713
Qy	719	PEPPEPNKSPPLTUSKEBSGWEPLTUSLEPAPARNPSSPERKATVPEQLOLEIELF	778
Db	714	PEPPEKAPALNISKEBSGTWEPLTSSLDPATAGPSPDDRATLPELELOLEIELF	773
Qy	779	LNSLSQFSLSEEQIILCSLISLSDSDDSEKNPSPKASQSSRTLSSGTVHWSQAEAR	838
Db	774	LNSLSQFSLSEEQIILCSLISLSDSDDSEKNPSPKASQSSRTLSSGTVHWSQAEAR	833
Qy	839	SSSMNMYLARGRPDTTFSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIATGSISSQIPAA	898
Db	834	TCSCSTALARGRPDTTFSYFNGVKVQIQSLNGEHLHIREFHRVKVGDIATGSISSQIPATA	893
Qy	899	FSLVTKDQGPVYDMEVPDSDGLDQCTLADPGSFAWSRWKVGQLENRP	947
Db	894	FSLVTKDQGPVYDMEVPDSDGLDQCTLADPGSFAWTVKVGQLENRP	942

DE	Hypothetical protein FLJ16568.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Esophageal;
RA	Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA	Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA	Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA	Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA	Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA	Wagatsuma M., Itakahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA	Z Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL	Submitted (MAR-2004) to the ENBL/GenBank/DBJ databases.
DR	EMBL; AK131438; BAD18584.1;
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0004672; P:protein kinase activity; IEA.
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR	InterPro; IPR011009; Kinase like.
DR	InterPro; IPR000719; Prot_kinase.
DR	Pfam; PF00069; Pkinase; 1.
DR	ProDom; PD000001; Prot_kinase; 1.
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW	Kinase.
SQ	SEQUENCE 477 AA; 52034 MW; ABBC068B4244B69E CRC64;
Query Match	45.5%; Score 2301; DB 2; Length 477;
Best Local Similarity	99.3%; Pred. No. 1.le-105;
Matches 431; Conservative	1; Mismatches 2; Indels 0; Gaps 0
QY	514 GDVADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDYIRGTTHMAPEVVLRGSC 573
Db	44 GGLAADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLLTGDYIRGTTHMAPEVVLRGSC 103
QY	574 DAKVDVWSSCCMLHLMNGCHPWTFQFRGPLCKIASEPPPVRIIPPCAPLTAQAIOEG 633
Db	104 DAKVDVWSSCCMLHLMNGCHPWTFQFRGPLCKIASEPPPVRIIPPCAPLTAQAIOEG 163
QY	634 LRKEPIHRVSAELGGKVNRALQOVGIKSPWRGEYKEPHRPPNQNYHOTLHAQPREL 693
Db	164 LRKEPIHRVSAELGGKVNRALQOVGIKSPWRGEYKEPHRPPNQNYHOTLHAQPREL 223
QY	694 SPRAGPRPAEITGRAPKLQPLPPPPPNNPKSPPLTLKSESGMWEPLPSLEPAPA 753
Db	224 SPRAGPRPAEITGRAPKLQPLPPPPPNNPKSPPLTLKSESGMWEPLPSLEPAPA 283
QY	754 RNSSPERKATVPQEHLQQLEIEILFNLSQLPSLEEQOILSCLSDLSDDSEKNP 813
Db	284 RNSSPERKATVPQEHLQQLEIEILFNLSQLPSLEEQOILSCLSDLSDDSEKNP 343
QY	814 SKASQSRRDTLSSGVHSWSACAARSSWNWVLARGPTPTPSYFNKGVKVQIQSLNGEHL 873
Db	344 SKASQSRRDTLSSGVHSWSQAARSSWNWVLARGPTPTPSYFNKGVKVQIQSLNGEHL 403
QY	874 HIREFHRVKVGDIATGISSQIPAAAFSLVTKGQPVRYDMVEVPDSGDIDLOCTLAPDGSA 933
Db	404 HIREFHRVKVGDIATGISSQIPAAAFSLVTKDQPVRYDMVEVPDSGDIDLOCTLAPDGSA 463
QY	934 WSWRVKHGLENRP 947
Db	464 WSWRVKHGLENRP 477
RESULT 5	
ID Q68D39	PRELIMINARY; PRT; 433 AA.
AC Q68D39;	
DT 25-OCT-2004	(TrEMBLrel. 28, Created)
DT 25-OCT-2004	(TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004	(TrEMBLrel. 28, Last annotation update)
DE	Hypothetical protein DKFPz686J04131 (Fragment).

RESULT 5	
Q68D39	
ID	Q68D39
AC	PRELIMINARY;
DT	PT; 433 AA.
DT	
DT	Q68D39;
DT	(TREMblrel. 28, Created)
DT	(TREMblrel. 28, Last sequence update)
DT	(TREMblrel. 28, Last annotation update)
DE	Hypotheical protein DKFZp686J04131 (Fragment).





Db 240 GLDFLHKKVHHDKPSNIVFMS--TKAVLVDFGLSVQMTED-----VTFPKDLR 288

Qy 558 GTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLGCHPWTOFF---RGPLCLKIA-SEPP 613

Db 289 GTTIYMSPEVILCRGHSTKADISLGATLIHMOTGTPDWVKRYPRSAVPSYLYIIHKQAP 348

Qy 614 PVREIPSCAPLTAQAIOEGLRKEPIHRVSAEELGGKVNRLAQVGGGLKSPWRGEYKEPR 673

Db 349 PLEDIADDCSPGRELIEASLERPNHRPRAADL-----LK----- 384

Qy 674 HPPPNQANYHOTLAQPRELSPRAPGPPPAEETGRAPKLQPLPPEPPEPKSPPLTSL 733

Db 385 -----HEALN-PPREDQPR-----QSLDSALLERKRLLSRLELELPENIADSSCTGS 431

Qy 734 KEESGM 739

Db 432 TESEEM 437

RESULT 8

M3K8 MOUSE

ID M3K8 MOUSE STANDARD; PRT; 467 AA.

AC Q07174;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Mitogen-activated protein kinase kinase 8 (EC 2.7.1.37) (COT

DE proto-oncogene serine/threonine-protein kinase) (C-COT) (Cancer Osaka

DE thyroid oncogene).

GN Name=Map3k8; Synonyms=Cot;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6 X CBA; TISSUE=Thymus;

RX MEDLINE=93308016; PubMed=8320169;

RA Ohara R., Miyoshi J., Aoki M., Toyoshima K.;

RT "The murine cot proto-oncogene: genome structure and tissue-specific

RT expression."

RL Jpn. J. Cancer Res. 84:518-525(1993).

CC -!- FUNCTION: Able to activate NF-kappa-B 1 by stimulating proteasome-

CC mediated proteolysis of NF-kappa-B 1/p105. Plays a role in the

CC cell cycle (By similarity).

CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.

CC -!- SUBUNIT: Interacts with NFkB1/p105.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- TISSUE SPECIFICITY: High in adult submandibular gland, thymus,

CC spleen and newborn digestive tract.

CC -!- PTM: Autophosphorylated (By similarity).

CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP

CC kinase kinase subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.iesb-sib.ch/announce/>

CC or send an email to [license@iesb-sib.ch](mailto:license@iesb-sib.ch)).

CC -----

CC EMBL; D13759; BRA02905.1; -.

DR PIR; I49609; I49609.

DR MGI; MGI:1346878; Map3k8.

DR InterPro; IPR011009; Kinase like.

DR InterPro; IPR000719; Prot kinase.

DR InterPro; IPR008271; Ser thr\_pkin\_AS.

DR Pfam; PF00069; Pkinase; I.

DR ProDom; PD000001; Prot kinase; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.

DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

KW ATP-binding; Phosphorylation; Proto-oncogene;

KW Serine/threonine-protein kinase; Transferase.

FT DOMAIN 138 388 Protein kinase.

FT NP BIND 144 152 ATP (By similarity).

FT BINDING 167 167 ATP (By similarity).

FT ACT SITE 253 253 Proton acceptor (By similarity).

SQ SEQUENCE 467 AA; 52941 MW; 60C2A34B530866BE CRC64;

Query Match 7.1%; Score 357.5; DB 1; Length 467;

Best Local Similarity 29.2%; Pred. No. 6.2e-10;

Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

Qy 379 GVLTTEKLKPDVYREER-----VHWATHQLRLG-----RGSGFEVHRMEDKQTGFQCAV 428

Db 107 GILLNMVISPNQGRYQIDSDVLLVPWKLTYRNIGSGFVPRGAFGVYLAQDMKTKRMAC 166

Qy 429 KKVLEVFRAEELMACAGLTSPIRVLYGVAREGPNVNIWFMELEGGSLGQLVKEOGLP 488

Db 167 KLIPIDQKPSDVEIQACFRHENIAELYGVLWGDTHVLFMEAGEGGSVLEKLESCGPR 226

Qy 489 EDRALYYLQGALEGLYHRRILHGDVADNVLLSSDGHAAALCDFGHAYCLQPDGLGK 548

Db 227 EFEEIIVTKHLKGLDFLHKKVHHDKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280

Qy 549 SLITGDYIP-----GTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLGCHPWTOFF---R 601

Db 281 -----VYLPKDLRGTEIYMSPEVILCRGHSTKADISLGATLIHMOTGTPDWVKRYPRSA 335

Qy 602 GPLCLKIA-SEPPPVREIPSCAPLTAQAIOEGLRKEPIHRVSAEELGGKVNRLAQVGG 660

Db 336 YPSVLYIIHKQAPPLEDIADDCSPGRELIEASLERPNHRPRAADL----- 382

Qy 661 LKSPWRGEYKEPRHPPPNQANYHOTLAQPRELSPR 696

Db 383 LK-----HEALN-PPREDQPR 397

RESULT 9

M3K8 RAT

ID M3K8 RAT STANDARD; PRT; 467 AA.

AC Q63552;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Mitogen-activated protein kinase kinase 8 (EC 2.7.1.37) (Tumor

DE progression locus 2) (TPL-2).

GN Name=Map3k8; Synonyms=Tpl2;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Long Evans; TISSUE=Liver;

RX MEDLINE=93211939; PubMed=7681591;

RA Patriotis C., Makris A., Bear S.E., Taichlis P.N.;

RT "Tumor progression locus 2 (Tpl-2) encodes a protein kinase involved

RT in the progression of rodent T-cell lymphomas and in T-cell

RT activation."

RL Proc. Natl. Acad. Sci. U.S.A. 90:2251-2255 (1993).

RN [2]

RP FUNCTION, AND MAP3K8-NFKB1/P105 BINDING.

RX MEDLINE=99133929; PubMed=9950430; DOI=10.1038/16946;

RA Belich M.P., Salmerton A., Johnston L.H., Ley S.C.;

RT "TPL-2 kinase regulates the proteolysis of the NF-kappaB-inhibitory

RT protein NF-kappaB1 p105."

RL Nature 397:363-368(1999).

CC -!- FUNCTION: Able to activate NF-kappa-B 1 by stimulating proteasome-

CC mediated proteolysis of NF-kappa-B 1/p105. Plays a role in the

CC cell cycle.

CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.

CC -!- SUBUNIT: Interacts with NFKB1/p105.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.



CC -!- TISSUE SPECIFICITY: Expressed in spleen, thymus, liver and lung.  
 CC -!- PTM: Autophosphorylated.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. MAP  
 CC kinase kinase subfamily.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M94454; AAA42185.1; --  
 CC PIR; A47388; A47388.  
 CC RGD; 620969; Map3k8.  
 CC InterPro; IPR011009; Kinase like.  
 CC InterPro; IPR007019; Prot\_kinase.  
 CC InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 CC Pfam; PF00069; Pkinase; 1.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC PROSITE; PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
 CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 CC ATP-binding; Phosphorylation; Proto-oncogene;  
 CC Serine/threonine-protein kinase; Transferase.  
 CC DOMAIN 146 388 Protein kinase.  
 CC NP\_BIND 144 152 ATP (By similarity).  
 CC BINDING 167 167 ATP (By similarity).  
 CC ACT\_SITE 253 253 Proton acceptor (By similarity).  
 CC SEQUENCE 467 AA; 52807 MW; 454E0E32768A4BD CRC64;

Query Match 7.1%; Score 356.5; DB 1; Length 467;  
 Best Local Similarity 29.2%; Pred. No. 6.9e-10;  
 Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY 379 GVLLTEKLPVDYVREE-----VHWATHQLRLG-----RGSFGEVHRMEDKQTGFQCAV 428  
 Db 107 GILLANVISFQNGRYQIDSDVLLVFWKLYTRSIGSGFVRGAFGVYLAQDMKTKRMAC 166  
 QY 429 KKVLEVFRAELMACAGLTSPIRVLYGAVREGPWNVLFMELGGSLGQLVKQGCLP 488  
 Db 167 KLIPVDQFSPDVEVQACFRHENIAELYGAVLVWGDVHLFMAGSGVLEKLESCGPMR 226  
 QY 489 EDRLAYLQALLEGLEYLHRSRILHGDVKADNVLSSDGSAAALCDFGHAVCLQPDGLGK 548  
 Db 227 EPEIIVVTKHVLKGLDFLHKKVHHDIKPSNIVFMS--TKAVLVDFGLSVQMTED---- 280  
 QY 549 SLLTGDYIP---GTETHMAPEVILGRSCDAKVDVWSSCCMLHMLNGCHPWTFP---R 601  
 Db 281 -----VYLPKDLRGTEIYMSPEVILCRGHSTKADITYSLGATLIHMQTGTPPWVKYPRSA 335  
 QY 602 GPLCLKIA-SEPPPVREITPPSCAPLTAQAIOSGLRKEPIHRVSAELGGKYNRALQQVGG 660  
 Db 336 YPSLYLIHKQAPPLEDAGDCSPGWRLEIAELERNHPRKPAADL----- 382  
 QY 661 LKSPWRGEYKEPRHPPPNQANYHTLHAQPRELSR 696  
 Db 383 LK-----HEALN-PPREDQPR 397

RESULT 10  
 Q3WTR2  
 ID Q3WTR2 PRELIMINARY; PRT; 1291 AA.  
 AC Q3WTR2;  
 DT 01-NOV-1999 (TremBLrel. 12, Created)  
 DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)  
 DE Apoptosis signal-regulating kinase 2.  
 GN Name=Map3k6;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Saitoh M., Ichijo H.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.  
 DR EMBL; AB021861; BAA78532.2; --  
 DR HSSP; Q16539; IKVI.  
 DR MGD; MGI:1855691; Map3k6.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; P:transferase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR011009; Kinase like.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 1291 AA; 143134 MW; 08DCE9D69D05E35 CRC64;  
 Query Match 7.0%; Score 356; DB 2; Length 1291;  
 Best Local Similarity 22.4%; Pred. No. 2.3e-09;  
 Matches 207; Conservative 127; Mismatches 350; Indels 240; Gaps 37;  
 QY 149 KSSSLAHAGVALAKPLRTPQEQSCTIPVQ---EDSPGAPVYRNTPTQTKLKEPGL 205  
 Db 411 EDSSELRLIWMKLAACLLAR-----KGVKEMQYVDVGVGLAQLANDP----- 455  
 QY 206 GQLCFKQGLGELRALPRSELHKLISPLQCLNHVWK-----LH-----PDGGPPLPPTH 257  
 Db 456 -----IQVLAALQYKLNAPWIYLVSVWETFLLYQHFRPTPEPSGGPLRAHF 504  
 QY 258 FPYRLPHPPFPFLQWKHPHLESLFKLACVDSQKPLDPDLHLSKLACVDSKPLGPGPH 317  
 Db 505 WLH-----FLQSCQPFKAPPOEDQCLVLVLEINKVLLPARLE-----IQGTDPMGAVT 554  
 QY 318 LEPSCLSRGAHEKFSVEYLVAHQGVSSQ-----AHSLSLAKTAAARSRSR-- 368  
 Db 555 L--SLEPETQEDSSWTFPTVTSICG-ISTSKLDQRCFLYALPPAQDVQLCFSPVERCQ 611  
 QY 369 -----BPSKTEDENEGVLLTEKLPVDYBYREVEVHWATHQLRLGRSGFGEVH 415  
 Db 612 RFCGLIQVLVNMNPDSSAPTEAEAG---AREVLEFDYSE---TGERLVLRGTGYGVVY 664  
 QY 416 RMEDKQTFQCAVKV-----RLEVFRAELMACAGLTSPIRVLYGAVREGPWNVIFME 470  
 Db 665 AGDRHTRVRIAIKEIPERSRFSQPLHEETIALHKLHKNIVRYLGSAGQGYLKIFME 724  
 QY 471 LLEGSSLGQLVKE--QGCLPEDRAL--YVLQALLEGLEYLHRSRILHGDVKADNVLSSDG 527  
 Db 725 EVPGSSLSLLRSVMGFLKDNESTISFYTRQLQGLSYLHENRIVHRDIKGNVLNIFS 784  
 QY 528 SHAALCDFGHAVCLQPDGLKSLITG-----DIYPGTETHMAPEV--LGRSCDAKVDVW 580  
 Db 785 GLLKISDFGTS-----KRLAGITPTCTFTGTQYMAPEIIDQGRGYKAAADIW 834  
 QY 581 SSCMMLHMLNGCHPWTFQFRGPLCLKIASEPPPVREI-----PSCAPLTAQA---IQ 631  
 Db 835 SLGCTVIEMATGPPHFE-----LGSFOAMFQVMGVKHPVPVPGSLSAEQAFL 885  
 QY 632 EGLRKEPIHRVSAELGGKYNRALQQVGLKSPWRGEYKEPRHP-----PNQANYH 683  
 Db 886 RTPEPDLRLASAEQLLG--DPFLQGRKRSR-----GSPRHTPRSGAPSGSSPSAD 938  
 QY 684 QTLHAQ---PRELSPRAGPRPAEE--TTGRAPKLQPLPP---EPPEPNKSPPLTTSKEE 736  
 Db 939 SATQSTQTFPRQAPSPQHPSPPKRCLUSYGDTSQLRVPEEPAEAPASPESSGLSLHQE 998





FT MUTAGEN 1381 1381 T->S: Reduced kinase activity and autophosphorylation.  
 FT MUTAGEN 1381 1381 T->E: Loss of kinase activity and autophosphorylation.  
 FT MUTAGEN 1381 1381 T->A: Falls to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.  
 FT MUTAGEN 1393 1393 T->S: Reduced kinase activity and autophosphorylation.  
 FT MUTAGEN 1393 1393 T->E: Loss of kinase activity and autophosphorylation.  
 FT MUTAGEN 1393 1393 T->A: Loss of kinase activity and autophosphorylation. Fails to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.  
 FT MUTAGEN 1394 1394 I->A: Loss of NF-kappa-B transcription factor activity and reduced ability to activate MAP2K1, MAP2K4, MAP2K7. No effect on AP-1 activity or activation of CHUK and IKKB. Loss of binding to IKKB.  
 FT MUTAGEN 1396 1396 F->A: Loss of AP-1 and NF-kappa-B transcription factor activity. Reduced ability to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.  
 FT MUTAGEN 1397 1397 M->A: Loss of AP-1 and NF-kappa-B transcription factor activity. Reduced ability to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.  
 FT MUTAGEN 1401 1401 V->A: Loss of AP-1 and NF-kappa-B transcription factor activity. Reduced ability to activate MAP2K1, MAP2K4, MAP2K7, CHUK and IKKB.  
 FT MUTAGEN 1402 1402 L->A: Loss of AP-1 transcription factor activity and reduced ability to activate CHUK and IKKB. No effect on NF-kappa-B activity or activation of MAP2K1, MAP2K4, MAP2K7. Loss of binding to MAP2K4.  
 FT MUTAGEN 1403 1403 R->A: Loss of AP-1 transcription factor activity. No effect on NF-kappa-B activity.  
 FT MUTAGEN 1404 1404 G->A: Loss of AP-1 and NF-kappa-B transcription factor activity.  
 FT CONFLICT 30 39 GGALQGSGA -> ALOQSG (in Ref. 2).  
 FT CONFLICT 103 103 Missing (in Ref. 2).  
 FT CONFLICT 257 257 V -> E (in Ref. 2).  
 FT CONFLICT 307 307 M -> V (in Ref. 2).  
 FT CONFLICT 413 413 S -> C (in Ref. 2).  
 FT CONFLICT 559 559 V -> A (in Ref. 2).  
 FT CONFLICT 883 883 V -> L (in Ref. 3).  
 FT CONFLICT 1467 1467 V -> L (in Ref. 3).  
 FT SEQUENCE 1493 AA; 161288 MW; CA65C9B7703C6BF9 CRC64;  
 Query Match 7.0%; Score 355.5; DB 1; Length 1493;  
 Best Local Similarity 23.6%; Pred. No. 2.9e-09;  
 Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;  
 QY 77 AISITQAECENSQSFSTFSEIRIFAGSKQVSQESLDQIPNNVAHATEGMARVCWK 136  
 DB 879 SLQAVAPTCSLEHVTREKTKGLSATRLSASEDISDLRAGVSGVLPSTTTEQ 938  
 QY 137 KR---RSKARKRKXKSKSLAHAGVALAKPLPRTPPEQESCTIPVQEDSPGAPYVNT 193  
 DB 939 PKPAVQTGRPHSQCLNSPLSHA---QLMFAPSPAPCSAPSV---DIS-----KHR 986  
 QY 194 PQFTPLKEPGLGQLCFKQLEGURPALPRSELHKLISPLQ---CLNHVMKLLHHPQDGGPL 251  
 DB 987 PQAFVPEKIPS-----ASPTQ-RKFSLQFQRCSEH-----RDSQDL 1023  
 QY 252 PLPTHFPYSLRHPFPHPPLQPKPHPLESFLGLKCLACVDSQKPLPDPLHSLKLA----- 305  
 DB 1024 ---SPVFTQSRPPPSNTHRPKPSRPV-----GSTSKLGDATKSSMTLDLGS 1069  
 QY 306 ---CVDS-----PKPLPGPH-----LEPSC-LSRGAHEKFSVEEVLVHALQGSVSS-- 347  
 DB 1070 SRCDDSGGGGNGNAVIPSDTFTVPEDKRLDVTNELSSIEDLLEASMPSSDTTIVT 1129

QY 348 -----SQASHLTSIAKTWAARSR---EPSP 372  
 DB 1130 FKSEVAVLSPEKAENDTYKDDVNNHQNCKEKWEAESEEAALAIAMAMSASODALPIVPOL 1189  
 QY 373 KTEDNEGVLITKLP-----VDYEVREHVHATHQLRLGRSGFSGVHRMEDKQTGF 424  
 DB 1190 QVNGEDIIIOQDTPELPGHTKAKQPYREDAEWLKGQ-QIGLGAFSSCYQADVGTGT 1248  
 QY 425 QCAVKV-----RLEVFA--EELMACAGLTPRIVPLYGAVREGPWNIIFELLE 473  
 DB 1249 LMAVQVTVRNTSSEQEEVEALREERIMGMHLNPNILMIGATCEKSNYNLFIEWMA 1308  
 QY 474 GSGSLQVKEQCLPEDRALYVLGOALSGLEYLHRRILHGDVADKADNVLLSSDGGHAALC 533  
 DB 1309 GGSVAHLISKYGAFKESVVIINYTEQLRLGLSYLHENQIHRDVKGANLLIDSTGQRLRIA 1368  
 QY 534 DFGHAVCLQPDGLGKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGC 593  
 DB 1369 DEGAARLASKGTGAGEFGQLL-GTIAFMAPEVLGRGOYGRSCDVWSVGCATIEMACAK 1427  
 QY 594 HPWTOFFRPG---PLCLKTIASE-----PPPVREIPSPCAPLTAQ--AIQEGRLREP 638  
 DB 1428 PPWNAEKHSNHLALIFKIASATTAPSIHSLSPGLRDVAVRCLELQPDQDRPPSRELLKHP 1487  
 QY 639 IHRVS 643  
 DB 1488 VFRIT 1492  
 RESULT 12  
 M3K2 MOUSE  
 ID\_M3K2 MOUSE STANDARD; PRT; 619 AA.  
 AC Q61083;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 2 (EC 2.7.1.37)  
 DE (MAPK/ERK kinase 2) (MEK kinase 2) (MEK 2).  
 GN Name=Map3K2; Synonyms=Mekk2;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96214986; PubMed=8621389; DOI=10.1074/jbc.271.10.5361;  
 RA Blank J.L., Gerwins P., Elliott E.M., Sather S., Johnson G.L.;  
 RT "Molecular cloning of mitogen-activated protein/ERK kinase kinases  
 (MEKK) 2 and 3. Regulation of sequential phosphorylation pathways  
 involving mitogen-activated protein kinase and c-Jun kinase.";  
 RL J. Biol. Chem. 271:5361-5368(1996).  
 RN [2]  
 RP FUNCTION, AUTOPHOSPHORYLATION, AND MUTAGENESIS OF PHE-571; ISO-573;  
 RP THR-575; GLN-576; PRO-577; PRO-580; LEU-582; PRO-583 AND VAL-586.  
 RX PubMed=12659851;  
 RA Huang J., Tu Z., Lee F.S.;  
 RT "Mutations in protein kinase subdomain X differentially affect MEKK2  
 and MEKK1 activity.";  
 RL Biochem. Biophys. Res. Commun. 303:532-540(2003).  
 RN [3]  
 RP INTERACTION WITH MAP2K5.  
 RX PubMed=12912994;  
 RA Nakamura K., Johnson G.L.;  
 RT "Pb1 domains of MEKK2 and MEKK3 interact with the MEK5 Pb1 domain for  
 activation of the ERK5 pathway.";  
 RL J. Biol. Chem. 278:36989-36992(2003).  
 CC -!- FUNCTION: Component of a protein kinase signal transduction  
 CC cascade. Regulates the JNK and ERK5 pathways by phosphorylating  
 CC and activating MAP2K5 and MAP2K7.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- COFACTOR: Magnesium.





Db	541	SPSSDDKSH	549	
RESULT 15				
Q9NYK3		PRELIMINARY;	PRT;	619 AA.
ID	Q9NYK3			
AC	Q9NYK3			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DE	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Protein kinase MEKK2b.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Wang C., Lo H.;			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF237978; AAF63496.1; -.			
DR	HSSP; Q13153; 1P3M.			
DR	GO; GO:0005524; F-ATP binding; IEA.			
DR	GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.			
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.			
DR	InterPro; IPR011009; Kinase like.			
DR	InterPro; IPR000270; OPR_PBI.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR002290; Ser_thr_kinase.			
DR	Pfam; PF00584; PBI; 1.			
DR	Pfam; PF00069; Pkinase; 1.			
DR	ProDom; PD000001; Prot_kinase; 1.			
DR	SMART; SM00666; PBI; 1.			
DR	SMART; SM00220; S_TKc; 1.			
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.			
KW	ATP-binding; Kinase.			
SK	SEQUENCE 619 AA; 69766 MW; 242BOF562D797873 CRC64;			
Query Match 6.9%; Score 350.5; DB 2; Length 619;				
Best Local Similarity 26.5%; Pred No. 1.9e-09;				
Matches 126; Conservative 64; Mismatches 180; Indels 105; Gaps 18				
QY	221	LPRLSELHLKLIPLQCLNHVWKLHPQDGG-----PLPLTHPPFPYGRPLPHFPFPHLPQ	273	
Db	193	IPES-MEQMLDPL-----SLSPENSGSGSCPSLDSPLDGESYKSRMPR-----	236	
QY	274	PWKPHLPESFLGKLACVDSQKPLDPHLSKLACVDSQKPLPGPHLEPSCLSRGAHEKRSV	333	
Db	237	-----AQSYPDNHQE---FSDYDNPIFEKFGKGTYPRRYHVSVHH	274	
QY	334	EY-----LVHALGQSVSSQA-----HSLSLAKTWAAGRSRRESPKTED---	376	
Db	275	QYQNGRKTTPRARRTQGNLTSPVSPSTHSLST-----SSGSIITPEYDDSR	326	
QY	377	-----NEGVLITKLPDYREYREHVHATHQLRLGRGSFGEVHRMEDKQTGPQCAVK	429	
Db	327	RRRGSIDNFTLTWDISPSRSPRAETNWLGL-LQGAAGRYLYCYDVTGRELAVK	385	
QY	430	KVLEVFRAE-----ELMACAGLTSPRIVPLYGAVR--RGPWNIFMELLEGSL	477	
Db	386	QVQFPDPSPETSKEVNALECEIQLLKNLLHERIVQYIGCLRDPQEKLTLSIFMEYMPGGS	445	
QY	478	GQLVKGGCLPEDRALYLLQALGLELYLHSRRILHGDVKADNVLLSSDGSAAALCDG-	536	
Db	446	KDQLKAYGALTENVTRYKTRQILEGVHYLHNSMTVHRDIKANTILRDTSTG-NVLGDFGA	504	
QY	537	---HAVCLQPDGLGKSLLTGDIYPGTETHMAPEVWLGRSCDAKDVWVSSCCMLHMLNG	592	
Db	505	SKELQITCLSGTM-KS-----VTGTPYWMSPEVISGGYGRKADIVSWACTVYEMLTE	557	
QY	593	CHPWTFRRGFLCLKIASEPPVREIPPSCAPLTAQAIQEGLRKEPIHRVSAAL	647	

Db 558 KPPWAEFEAMAAIFKIAIOTNP-K-LPPHVSDYTRDFLKR-IFVEAKLRPSADEL 610

Search completed: June 7, 2005, 12:02:21  
Job time : 143.585 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2005, 11:43:10 ; Search time 32.7215 Seconds  
(without alignments)  
2784.626 Million cell updates/sec

Title: US-09-155-676B-7  
Perfect score: 5052  
Sequence: 1 MAVNEMACPGAGSAGCQK.....PDGSFAWSRVKHCQLENRP 947

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	360	7.1	467	2 A48713	serine/threonine-s
2	357.5	7.1	467	2 I49609	proto-oncogene pro
3	356.5	7.1	467	2 A47388	serine/threonine p
4	355.5	7.0	687	2 A46212	MEK kinase - mouse
5	353	7.0	1493	2 T10757	MAP kinase kinase
6	336	6.7	1288	2 JB0363	mitogen-activated
7	331	6.6	415	2 A41253	kinase-related tra
8	325	6.4	651	2 A96591	NPX1-related prote
9	325	6.4	1478	2 S20117	protein kinase BCK
10	319	6.3	653	2 T34356	hypothetical prote
11	313.5	6.2	883	2 A96662	hypothetical prote
12	307	6.1	696	2 S55694	protein kinase (EC
13	305.5	6.0	525	2 S58682	protein kinase, p2
14	305	6.0	550	2 T36746	probable serine/th
15	304	6.0	608	2 G96575	probable MEK kinas
16	303.5	6.0	1116	2 T38073	serine/threonine-p
17	301	6.0	706	2 T48084	STE11 protein kina
18	301	6.0	939	2 S28394	probable serine/th
19	298	5.9	1387	2 T16511	hypothetical prote
20	295	5.8	836	2 B96716	probable serine/th
21	294	5.8	1338	2 T30565	MAP kinase kinase
22	294	5.8	1401	2 T39225	MAP kinase kinase
23	292	5.8	544	2 S40482	serine/threonine-s
24	291.5	5.8	543	1 JC4070	protein kinase (EC
25	291.5	5.8	756	2 T50298	MAP kinase kinase
26	291	5.8	545	2 G01793	p21-activated prot
27	291	5.8	556	2 T36502	serine/threonine p
28	288.5	5.7	622	2 T15467	hypothetical prote
29	288	5.7	544	2 A57597	beta-p21-activated

30 287 5.7 1734 2 A54602 microtubule-associ  
31 284 5.6 372 2 T52621 mitogen-activated  
32 282 5.6 490 2 S47946 protein kinase hom  
33 282 5.6 1062 2 S46367 protein kinase CDC  
34 281 5.6 556 2 T42100 serine/threonine p  
35 280 5.5 544 2 I49376 p21 activated kina  
36 280 5.5 738 1 S51380 protein kinase STE  
37 279 5.5 1230 2 T18256 probable serine/th  
38 279 5.5 1230 2 T18259 serine/threonine p  
39 279 5.5 1370 2 T19188 hypothetical prote  
40 277.5 5.5 481 2 JE0377 p70 S6 kinase (EC  
41 276.5 5.5 363 2 T51735 mitogen-activated  
42 276.5 5.5 582 2 T51625 MAP3K alpha protei  
43 276.5 5.5 626 2 B70754 probable serine/th  
44 275 5.4 426 2 S71886 Ste20-like protein  
45 275 5.4 1579 2 S59801 protein kinase SSK

## ALIGNMENTS

## RESULT 1

A48713  
serine/threonine-specific protein kinase cot. 58K form - human  
N:Alternate names: cot proto-oncogene, 58K form  
N:Contains: serine/threonine-specific protein kinase cot, 52K form  
C:Species: Homo sapiens (man)  
C:Date: 07-Apr-1994 #sequence revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A48713; B48713; S31639  
R:Aoki, M.; Hamada, F.; Sugimoto, T.; Sumida, S.; Akiyama, T.; Toyoshima, K.  
J. Biol. Chem. 268, 22723-22732, 1993  
A:Title: The human cot proto-oncogene encodes two protein serine/threonine kinases with  
A:Reference number: A48713; MUID:94043034; PMID:8226782  
A:Accession: A48713  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-467 <AOK>  
A:Cross-references: UNIPROT:P41279  
A:Experimental source: TCO-4 cells  
A>Note: sequence inconsistent with nucleotide translation  
A>Note: sequence extracted from NCBI backbone (NCBIN:138970, NCBI:P138971)  
A:Accession: B48713  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 30-467 <A02>  
A:Experimental source: TCO-4 cells  
A>Note: sequence inconsistent with nucleotide translation  
A>Note: sequence extracted from NCBI backbone (NCBIN:138970, NCBI:P138972)  
R:Chan, A.M.; Chedid, M.; Aaronson, S.A.; Miki, T.; McGovern, E.S.  
submitted to the EMBL Data Library, July 1992  
A:Description: A transforming gene isolated by expression cloning from Ewing's sarcoma c

Query Match 7.1%; Score 360; DB 2; Length 467;

Best Local Similarity 27.2%; Pred. No. 1.5e-09;

Matches 116; Conservative 69; Mismatches 155; Indels 86; Gaps 14;

Qy 332 SVEYLVHVLQGSVSSQAHSLTSLAKTWAARGSRSPSPKTEQNEGVLTLEKLKPDY 391

Db 80 TVEDLL--AFANHISNTAKHY-----GORPOE-----SGILLNMWITPQNG 119

Qy 392 EYREE-----VHWATHQLRLQ-----RGSFGFVHRMEDKQTGFCVAKVKVLEVFRAEEL 441

Db 120 RYQIDSDVLLIPKWLTYRNTGSDFIPIRGAFKGYLAQDIKTKRMACKLIIPVDQFKPSDV 179

Qy 442 MACAGLTSPIRVPLYGAVREGSPWYNIPMELLEGSLGQLVKEGQCLPEDRALYYLQGALE 501

```
Db 180 EIQACFRHENTIAELYGAVLWGTETVHLFMEAGEGGSVLEKLESCGPRREFIIVWTKVLK 239
QY 502 GLEVLHRSRIILHGVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGKSLTGYIP----- 557
Db 240 GLDFLHKKIHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED-----VYFFKDLR 288
QY 558 GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF---RGPLCLKIA-SEPP 613
Db 289 GTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRAYPSYLIYIHKQAP 348
QY 614 PVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGGKVNRAALQOQVGLKSPMRGEYKEPR 673
Db 349 PLEDIADDCSPGMRELIEASLERNPNRPRAADL-----LK----- 384
QY 674 HPPPNQANYHOTLHAQPRELSPRAPGRPAEETTGAPKLOPLPPEPPKNSPLPLTUS 733
Db 385 -----HEALN-PPREDQPRC---QSLDSALLERKLLSRKELELPENTADSSCTGS 431
QY 734 KERSGM 739
Db 432 TESEEM 437

RESULT 2
149609
proto-oncogene protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49609
R;Ohara, R.; Miyoshi, J.; Aoki, M.; Toyoshima, K.
Jpn. J. Cancer Res. 84, 518-525, 1993
A;Title: The murine cot proto-oncogene 1: genome structure and tissue-specific expression
A;Reference number: I49609; MUID:93308016; PMID:8320169
A;Accession: I49609
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-467 <RES>
A;Cross-references: UNIPROT:Q07174; GB:D13759; NID:G286072; PIDN:BAA02905.1; PID:G286073
C;Genetic:
A;Gene: cot
F;136-388/Domain: protein kinase homology <KIN>

Query Match 7.1%; Score 357.5; DB 2; Length 467;
Best Local Similarity 29.2%; Pred. No. 1.9e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY 379 GVLLTEKLKPDVDEYREE-----VHWATHQLRLG-----RGSFGVHRMEDKOTGFQCAV 428
Db 107 GILLNNVISPQNGRYQIDSDVLLVPWKLTYRISGSPVPRGAFGKVYLAQDMKTKRWAC 166
QY 429 KKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGQLVKEQGCLP 488
Db 167 KLIPIDQFKPSDVEIQACFRHENIAELYGAVLWGTETVHLFMEAGEGGSVLEKLESCGPMR 226
QY 489 EDRALYILGQALEGLEVYLSRRIILHGVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGK 548
Db 227 EFEIIVWTKVHLKGLDFLHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280
QY 549 SLLTGDYIP---GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF---R 601
Db 281 -----VYLPKDLRGTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRSA 335
QY 602 GPLCLKIA-SEPPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGGKVNRAALQOQVGG 660
Db 336 YPSYLYIHKQAPPLEDIAGDCSPGMRELIEAALERNPNRPKAADL----- 382
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPR 696
Db 383 LK-----HEALN-PPREDQPR 397

RESULT 3
149609
proto-oncogene protein - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49609
R;Ohara, R.; Miyoshi, J.; Aoki, M.; Toyoshima, K.
Jpn. J. Cancer Res. 84, 518-525, 1993
A;Title: The murine cot proto-oncogene 1: genome structure and tissue-specific expression
A;Reference number: I49609; MUID:93308016; PMID:8320169
A;Accession: I49609
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-467 <RES>
A;Cross-references: UNIPROT:Q07174; GB:D13759; NID:G286072; PIDN:BAA02905.1; PID:G286073
C;Genetic:
A;Gene: cot
F;136-388/Domain: protein kinase homology <KIN>

Query Match 7.1%; Score 357.5; DB 2; Length 467;
Best Local Similarity 29.2%; Pred. No. 1.9e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY 379 GVLLTEKLKPDVDEYREE-----VHWATHQLRLG-----RGSFGVHRMEDKOTGFQCAV 428
Db 107 GILLNNVISPQNGRYQIDSDVLLVPWKLTYRISGSPVPRGAFGKVYLAQDMKTKRWAC 166
QY 429 KKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGQLVKEQGCLP 488
Db 167 KLIPIDQFKPSDVEIQACFRHENIAELYGAVLWGTETVHLFMEAGEGGSVLEKLESCGPMR 226
QY 489 EDRALYILGQALEGLEVYLSRRIILHGVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGK 548
Db 227 EFEIIVWTKVHLKGLDFLHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280
QY 549 SLLTGDYIP---GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF---R 601
Db 281 -----VYLPKDLRGTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRSA 335
QY 602 GPLCLKIA-SEPPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGGKVNRAALQOQVGG 660
Db 336 YPSYLYIHKQAPPLEDIAGDCSPGMRELIEAALERNPNRPKAADL----- 382
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPR 696
Db 383 LK-----HEALN-PPREDQPR 397
```

```
A47388
serine/threonine protein kinase - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A47388
R;Patriotis, C.; Makris, A.; Bear, S.E.; Tsichlis, P.N.
Proc. Natl. Acad. Sci. U.S.A. 90, 2251-2255, 1993
A;Title: Tumor progression locus 2 (Tpi-2) encodes a protein kinase involved in the pro
A;Reference number: A47388; MUID:93211939; PMID:7681591
A;Accession: A47388
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-467 <PAP>
A;Cross-references: UNIPROT:Q63562; GB:M94454; NID:G207082; PIDN:AAA42185.1; PID:G207083
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIN:128133, NCBI:P:128134)
F;136-388/Domain: protein kinase homology <KIN>

Query Match 7.1%; Score 356.5; DB 2; Length 467;
Best Local Similarity 29.2%; Pred. No. 2.1e-09;
Matches 98; Conservative 60; Mismatches 115; Indels 63; Gaps 10;

QY 379 GVLLTEKLKPDVDEYREE-----VHWATHQLRLG-----RGSFGVHRMEDKOTGFQCAV 428
Db 107 GILLNNVISPQNGRYQIDSDVLLVPWKLTYRISGSPVPRGAFGKVYLAQDMKTKRWAC 166
QY 429 KKVRLVFRABELMACAGLTSPRIVPLYGAVREGPWNIIFMELLEGGSLGQLVKEQGCLP 488
Db 167 KLIPVDQFKPSDVEIQACFRHENIAELYGAVLWGTETVHLFMEAGEGGSVLEKLESCGPMR 226
QY 489 EDRALYILGQALEGLEVYLSRRIILHGVKADNVLLSSDGSAAALCDFGHAVCLQPDGLGK 548
Db 227 EFEIIVWTKVHLKGLDFLHKKVHHIDIKPSNIVFMS--TKAVLVDFGLSVQMTED----- 280
QY 549 SLLTGDYIP---GTETHMAPEVVLGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF---R 601
Db 281 -----VYLPKDLRGTEIYMSPEVILCRGHTKADIYSLGATLIHMOTGTPPWKRYPRSA 335
QY 602 GPLCLKIA-SEPPPVREIPSPCAPLTAQAIQEGLRKEPIHRVSAALGGKVNRAALQOQVGG 660
Db 336 YPSYLYIHKQAPPLEDIAGDCSPGMRELIEAALERNPNRPKAADL----- 382
QY 661 LKSPWRGEYKEPRHPPPNQANYHOTLHAQPRELSPR 696
Db 383 LK-----HEALN-PPREDQPR 397

RESULT 4
A46212
MEK kinase - mouse
C:Species: Mus musculus (house mouse)
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 28-May-1999
C:Accession: A46212
R;Lange-Carter, C.A.; Pleiman, C.M.; Gardner, A.M.; Blumer, K.J.; Johnson, G.L.
Science 260, 315-319, 1993
A;Title: A divergence in the MAP kinase regulatory network defined by MEK kinase and Rat
A;Reference number: A46212; MUID:93227040; PMID:8395802
A;Accession: A46212
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-687 <LAN>
A;Experimental source: brain
A;Note: sequence extracted from NCBI backbone (NCBI:P:129292)
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C;Keywords: ATP
F;416-683/Domain: protein kinase homology <KIN>
F;424-432/Region: protein kinase ATP-binding motif

Query Match 7.0%; Score 355.5; DB 2; Length 687;
Best Local Similarity 23.6%; Pred. No. 3.5e-09;
Matches 157; Conservative 81; Mismatches 278; Indels 149; Gaps 24;

QY 77 AISIIAQECENSQSFSTFSEIRIFITAGSKQYSQSESLDQIPNNVAHATEGKMARVCWK 136
```





```
QY 149 KSSKSLAHAGVALAKPLPRTPQESCTIPQV---EDESPILGAVYVRNTPTQTKPLKEPGL 205
Db 410 EDSKELRLIGMKLCLLAR-----KGCVEKQYWDVGVYGAQILANDP-----454
QY 206 QOLCFKQOLGEGRLPALPRSELHLKILSPLOCLINHVK---LHHQDGGPLPLPHTPPYSR 262
Db 455 -----TQVLAAEOLYKLNAPIWLVSVMETFLY--QHGRPTPEP--PGGPPR 499
QY 263 LPHPPFPHPQPKPHEPFLGKLACVDSQKPLPDPHLSKLACVSPKPLPGPHLEPSC 322
Db 500 RAHFWLHFLQSQQPF-----KTACAQGDQCL-----VLVLEMKVLLPAKLE---542
QY 323 LSRGAHEKFSVEEYLVHALQSVSSQAHSLSLTKTWAAGSR-----366
Db 543 -VRGTDPSVTLSLLEPETQDIPSSWTFPVAISGVASAKRDERCCFLYALPPAODVOL 601
QY 367 -----SRPSPKTNEGVLLTEKLPVDYFYREVEVHWATHQLRL 406
Db 602 CFPVSVGHQWFCGLIQAWVTNPDSTAPAEAEAG---AEEMLEPDEYETE---TGERLVL 654
QY 407 GRGSFGFVHRMEDKQTFQCAVKKV-----RLEVFRAEELMACAGLTSPIRIVLYCAVRE 461
Db 655 GKTYGVVYAGRDHRTVRIKIEIPERSRFPQPIHEETALHRRHKNIVRYLGSASQ 714
QY 462 GPWNIPFMELEGSLGQVKE-QGCLPEDRAL--YYLQGALEGLLEYLSRRILHGDVKA 518
Db 715 GGYLKIFMEIEVPGCSLLSLRSVMGPKDNESTISFYTRQLQGLYLDHNDHIVHRDIKG 774
QY 519 DNVLLSSDGSAAALCDFGHAVCLQDGLGKSLITG-----DYIPGTETHMAPEVV--LGR 571
Db 775 DNVLIINTFGSLKISDFGTS-----KRLAGITPCTETPTGTLOYMAPEIIDQGR 824
QY 572 SCDKAVDVWSSCCMMLHMLNGCHPWTQFFRGPLCLKIASEPPPVREI-----BPSCAPL 625
Db 825 GYCKADIWSLGTVTIEMATGRPFHE-----LGSFQAMQVGMVKVHPWPMSL 875
QY 626 TAQA---IOEGLRKEPIHRVSAEELGGKVNRAALQVGGGLKSPWRGEYKEPRHPPPNQANY 682
Db 876 SAEAQAFLRLTRPEPDLRSASQTLTG--DPFLQPGKRSRP-----SSPRHAP-----922
QY 683 HQTLLHAQRELSPRAGPRPAETT-----GRAPKLOPLPPE-----PP 722
Db 923 -----RPDAPSASTP--SANSSTQSFPCQAPSOHPSPPKRCKSYGTSQLRVPE 975
QY 723 EPNKSPPLTSLKEESGMWEPLSLLEPAPARNPSPERKATVPBQELQQLIEFLNLSL 782
Db 976 EPAAEERPAS-PESSG-----LSLLHQESKRR-----AMLAVALBQELPALAENLHQEQK 1024
QY 783 SQP---FSLEEOEQILSCL 798
Db 1025 OEQGARLGRNHVLELLRCL 1043

RESULT 7
A41253
Kinase-related transforming protein (EC 2.7.1.-) - human
C:Species: Homo sapiens (man)
C:Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 15-Mar-2004
C:Accession: A41253
R:Miyoishi, J.; Higashi, T.; Mukai, H.; Ohuchi, T.; Kakunaga, T.
Mol. Cell. Biol. 11, 4088-4096, 1991
A:Title: Structure and transforming potential of the human cot oncogene encoding a putat
A:Reference number: A41253; PMID:91304400; PMID:2072910
A:Accession: A41253
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-415 <MIV>
C:Keywords: phosphotransferase
F:136-388/Domain: protein kinase homology <KIN>

Query Match
Best Local Similarity 6.6%; Score 331; DB 2; Length 415;
Matches 104; Conservative 65; Mismatches 136; Indels 70; Gaps 13;
```

```
QY 332 SVEEYLVHALQSVSSQAHSLSLTKTWAAGSRSPKPTKEDNEGVLITTEKLPVDY 391
Db 80 TVEDLL--AFANHISNTAKHFY-----GQRPE-----SGILLNMVITPQNG 119
QY 392 EYREE-----VHWATH---OLRLGRSGFGEVHRMEDKQTFQCAVKKVLEVFRAEEL 441
Db 120 RYQIDSDVLLIISNTAKHFYQRPQESGILLKLYLAQDIKTKRWACKLIPVDQFKPSDV 179
QY 442 MACAGLTSPIRIVLYCAVREGFWNNIFMELLEGGSLGOLVKEQGCCLPEDRALYYLQALE 501
Db 180 ETQACFRHENTAEILGAVLWGETVHLFMEAGEGSVLEKLSGCGPMRFEFIIWVTKHLK 239
QY 502 GLEYLHSRRILHGDVKAENVLLSSDPGSHAAALCDFGHAVCLQDGLGKSLITGDYIP----557
Db 240 GLDFLHSKKVIHDIKPSNIVFMS--TKAVLVDFGLSVQMTED-----VYFPKDLR 288
QY 558 GTETHMAPEVVILGRSCDAKVDVWSSCCMMLHMLNGCHPWTQFF---RGPLCLKIA-SEPP 613
Db 289 GTEIYMSPEVILCRGHSTKADIYSLGATLIHMGTGTPPVKRYPRPSAYPSYLYIIHKQAP 348
QY 614 PVREIPPSCAPLTAQAIQEGRLKEPIHRVSAEELGGKVNRAALQVGGGLKSPWRGEYKEPR 673
Db 349 PLEDIADDSCPMRELIIEASLERPNHRPRAUL-----LK-----HEAL 388
QY 674 HPP-PNQANYHQTLH 687
Db 389 NPREDQPRGHQVIH 403

RESULT 8
A96591
NPKI-related protein kinase 2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A96591
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; PMID:21016719; PMID:11130712
A:Accession: A96591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-651 <STO>
C:Cross-references: UNIPROT:Q9FZ36; GB:AE005173; NID:g9857521; PIDN:AAG00876.1; GSPDB:G
C:Genetics:
A:Gene: T24C10.7
A:Map position: 1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match
Best Local Similarity 6.4%; Score 325; DB 2; Length 651;
Matches 132; Conservative 69; Mismatches 188; Indels 116; Gaps 18;

QY 383 TEKLKPYDYEYRBEVHWATHQLRLGRSGFGEVHRMEDKQTFQCAVKKVRL-----EV 435
Db 59 TVQIKP-----PIRWKQQL-IGRGAFTVYMGWNLDSGELLAKVQLITSNCASKEK 110
QY 436 FRA-----BELMACAGLTSPIRIVLYCAVREGFWNNIFMELLEGGSLGOLVKEQGCCLPE 489
Db 111 TOAHIOELEEEVKLLKNLSHPNIVRYLGTVREDETILNILEFVPGGSISSLEKFGAPPE 170
QY 490 DRALYYLGOALEGLEYLHSRRILHGDVKAENVLLSSDGSAAALCDFG---HAVCLQPDGL 546
Db 171 SVVTTYNQLLGLEYLHSHAIHWRDIKGNILVDNQGCG-IKLADFGASKQVLAETISG 229
```

Qy 547 GKSLTGDYIPGTETHMAPEVVLGRSCDAKVDVWSSCCMLHMLNGCHPWTOFFRGPLCL 606  
Db 230 AKSM-----KGTPTWMAPEVILQTGHGFSADISVGVCTVIEMTVGKAPWSQQYK----- 278  
Qy 607 KIA-----SEPPVRETPPSCAPLTAQAIOEGLAKEPHRVSAAELGKYNRALQQ 657  
Db 279 EIAAIFHGTYSKHP-----IPDNISSDANFLLKCLQEBNLRPTASEL----- 324  
Qy 658 VGGLKSPW-RGEYKEPRHPPPNOANYHOTLAQPRELSFAPRGPRAEETTRAPKLQPP 716  
Db 325 ---LKHPPVTGQKE-----SASKDLTSFMDNSCSP 352  
Qy 717 LPPEPPENKSPPLTSLKEESGMWEPLLSLEAPARNPSPERKATVPQEOLOOLEIE 776  
Db 353 LPSE--LTNITSYQTSDDVG--DICNLGLTCTLA-----FPEKSIQNNSLC 397  
Qy 777 LFLNSLQPFSLQEOQLSCLSIDLSLSDSEKNPS-----KASQSRDRTLSSGVHWS 832  
Db 398 LKSNV-----GYDDDDNDMCLIDDENFLYNGETGPSLDNNTDAKSCDTMSEISDLK 452  
Qy 833 SQAEARSSWNVLARGRPTTPSY 857  
Db 453 CKFDENSGNETETKVSMEVDHPSY 477

RESULT 9  
S20117  
protein kinase BCK1 (BC 2.7.1.1) - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein J0906; protein kinase SLK1; protein kinase SSP31; protein YJL  
C:Species: Saccharomyces cerevisiae  
C:Date: 23-Apr-1993 #sequence revision 23-Apr-1993 #text\_change 16-Aug-2004  
C:Accession: S20117; S20298; S22285; S19061; JQ1432; S56872; S30794; JQ1118  
R:Costigan, C.; Gehring, S.; Snyder, M.  
Mol. Cell. Biol. 12, 1162-1178, 1992  
A:Title: A synthetic lethal screen identifies SLK1, a novel protein kinase homolog impli  
A:Reference number: S20117; MUID:92186847; PMID:1545797  
A:Accession: S20117  
A:Molecule type: DNA  
A:Residues: 1-1478 <COS>  
A:Cross-references: UNIPROT:Q01389; EMBL:M84389  
A:Experimental source: strain S288C  
R:Miosga, T.; Boles, E.; Schaff-Gerstenschlaeger, I.; Schmitt, S.; Zimmermann, F.K.  
Yeast 10, 1481-1488, 1994  
A:Title: Sequence and function analysis of a 9.74 kb fragment of Saccharomyces cerevisia  
A:Reference number: S50295; MUID:95176706; PMID:7871887  
A:Accession: S50298  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-1478 <MIO>  
A:Cross-references: EMBL:X77923; NID:9640004; PIDN:CAA54896.1; PID:9640009  
R:Lee, K.S.; Levin, D.E.  
Mol. Cell. Biol. 12, 172-182, 1992  
A:Title: Dominant mutations in a gene encoding a putative protein kinase (BCK1) bypass b  
A:Reference number: S22285; MUID:92107166; PMID:1729597  
A:Accession: S22285  
A:Molecule type: DNA  
A:Residues: 1-58, 'I', 60-1478 <LEE>  
A:Cross-references: EMBL:X60227  
A:Experimental source: strain EG123  
R:Lee, K.S.; Levin, D.E.  
submitted to the EMBL Data Library, June 1991  
A:Description: An extragenic suppressor of mutations in the S. cerevisiae protein kinase  
A:Reference number: S19061  
A:Accession: S19061  
A:Molecule type: DNA  
A:Residues: 1-58, 'I', 60-263, 'P', 265-278, 'I', 280-702, 'S', 707-708, 'KP', 714, 'VITWTE', 715-79  
A:Cross-references: EMBL:X60227; NID:G3414; PIDN:CAA42788.1; PID:G3415  
A:Experimental source: strain EG123  
R:Irie, K.; Araki, H.; Oshima, Y.  
Gene 108, 139-144, 1991  
A:Title: A new protein kinase, SSP31, modulating the SMP3 gene-product involved in plas  
A:Reference number: JQ1432; MUID:92104496; PMID:1840547  
A:Accession: JQ1432

A:Molecule type: DNA  
A:Residues: 149-1478 <IRI>  
A:Cross-references: EMBL:D10389; DDBJ:D90446  
R:Miosga, T.; Schaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chalwatzis, N.; Fournie  
submitted to the Protein Sequence Database, September 1995  
A:Reference number: S56855  
A:Accession: S56872  
A:Molecule type: DNA  
A:Residues: 1-1478 <MIW>  
A:Cross-references: EMBL:Z49370; NID:G1008269; PIDN:CAA89389.1; PID:G1008270; MIPS:YJL095w  
R:Cusick, M.E.  
submitted to the EMBL Data Library, March 1992  
A:Reference number: S27437  
A:Accession: S30794  
A:Molecule type: DNA  
A:Residues: 602-959, 'R', 961, 'R', 963-1085, 'V', 1087, 'SLLIHT', 1092-1094, 'RMD', 1101, 'TV', 11  
A:Cross-references: EMBL:M86604; NID:G172073; PIDN:AAA31179.1; PID:G172074  
C:Genetics:  
A:Gene: SGD:BCK1; SLK1; SSP31  
A:Cross-references: SGD:S0003631; MIPS:YJL095w  
A:Map position: 10L  
C:Function:  
A:Description: phosphotransferase; protein kinase; involved in cell proliferation  
C:Superfamily: protein kinase homology  
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein k  
F:1173-1440/Domain: protein kinase homology <KIN>  
F:1181-1189/Region: protein kinase ATP-binding motif

Query Match 6.4%; Score 325; DB 2; Length 1478;  
Best Local Similarity 28.7%; Pred. No. 1.9e-07;  
Matches 117; Conservative 53; Mismatches 163; Indels 74; Gaps 14;  
Qy 258 FPYRLPHPPFPFHPQLPWKHPLESFLGKACVDSQKPLPDPHLSKLACVDSPKPLPGPH 317  
Db 1043 FPRANLQKPIGTEGASTSPKSLDLSL-----SPKNVASSR 1078  
Qy 318 LEPSCLSRGAHEKFSVEYLVAHQGS---VSSQASHLSLTAKTMAARGSRSPSPK- 373  
Db 1079 TEPSTSPRPVPDSSY-EFIQDGLGNKKNPLQNKPKTKTKTIRTIHAESLARKNSVKL 1137  
Qy 374 -----TEDNEGVLITKLPVDYRYREVVHATHQLRLGRGSGEVHRME 418  
Db 1138 KRQNTKMGTRMVEVTNNH---MVSINKAKNSKGYK-EFAMKGM-IGKSGFQAVYLC 1193  
Qy 419 DKQTGFCAVKKVRLEVFRAE-----ELMACAGLTSPRIPLYGAVREGPWV 465  
Db 1194 NVTTGEMVAKQVEVPKYSSQNEALSTVEALREVSTLKOLDHLNIVQYLGFEKNKIY 1253  
Qy 466 NIFMELLEGGSLGQLVKEQGLPEDRALYYLQGALEGLEYLHSRRILHGDVKADNVLS 525  
Db 1254 SLFLEYVAGGSGVSLIRMYGRFDBPLIKHLTTQVLKGLAYLHSGILHRDMKADNLLDQ 1313  
Qy 526 DGSAAALCDRGHACVCLQPDGLGKSLLTGDIYPTGTHMAPVW-LGRSCDAKVDVWS 584  
Db 1314 DGI-CKISDFGIS-RKSKDIYSNDSMT---MRGTVFWMAPEMVMTKQGSYAKVDIWS 1368  
Qy 585 MMLHMLNGCHPWTOFFRGPLCLKIA---SEPPVRETPPSCAPLTAQ 628  
Db 1369 IVLEMPAGKRPWSNLEVVVAMFKIGKSKSAPP-----IPEDTLPLISQ 1411

RESULT 10  
T34356  
hypothetical protein T19A5.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T34356  
R:Bradshaw, H.  
submitted to the EMBL Data Library, March 1996  
A:Description: The sequence of C. elegans cosmid T19A5.  
A:Reference number: Z21512  
A:Accession: T34356  
A:Status: preliminary; translated from GB/EMBL/DBJ



Db 814 QDRSSRPTPHVFWNDNGSIQPCYNWN-----KDNQPVLSHDVSSQQLLS-----EHLK 863

Qy 875 IREFHRVKVGIATGISS 892

Db 864 LKSL-----DLRPGFST 875

RESULT 12

S55694

protein kinase (EC 2.7.1.37) sk1, CAMP-dependent - fission yeast (Schizosaccharomyces p

C;Species: Schizosaccharomyces pombe

C;Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 16-Aug-2004

C;Accession: S55694; T38040

R;Jin, M.; Fujita, M.; Culley, B.M.; Apolinario, E.; Yamamoto, M.; Maundrell, K.; Hoffma

Genetics 140, 457-467, 1995

A;Title: sk1, a high copy number suppressor of defects in the CAMP-dependent protein ki

A;Reference number: S55694; MUID:96120227; PMID:7498728

A;Accession: S55694

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-696 <JIN>

A;Cross-references: UNIPROT:P50530; GB:D38108; NID:g1136301; PIDN:BAA07286.1; PID:d10078

A;Note: the authors translated the codon GAT for residue 687 as His

R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Whitehead, S.; Churcher, C.M.

submitted to the EMBL Data Library, August 1999

A;Reference number: Z21764

A;Accession: T38040

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-176, 'F', 178-198, 'A', 200-696 <MCD>

A;Cross-references: EMBL:AL109951; PIDN:CAB53053.1; GSPDB:GN00066; SPDB:SPAC1B9.02c

A;Experimental source: strain 972h; cosmid c1B9

C;Genetics:

A;Gene: sk1; SPDB:SPAC1B9.02c

A;Map position: 1

A;Introns: 80/3; 311/1; 633/2

C;Superfamily: protein kinase homology

C;Keywords: ATP; phosphotransferase; protein kinase

F;300-563/Domain: protein kinase homology <KIN>

F;308-316/Region: protein kinase ATP-binding motif

Query Match 6.1%; Score 307; DB 2; Length 696;

Best Local Similarity 21.5%; Pred. No. 5.6e-07;

Matches 187; Conservative 125; Mismatches 292; Indels 264; Gaps 42;

Qy 36 QSSVYKLEAVEKSPVFCGKWEILNDVITKGTAKGSEAGPAAISIIAQACENSQEFSP 95

Db 19 QASPTIQSHSTQPV-----LSNDHSTKVNDYEGKEG-----SSNGYDP- 58

Qy 96 FSEIFTAGSKQYSQSESLDQIPNNVAHATEGKMARVCWKGKRSKARKKKKSKSLA 155

Db 59 ----VFMSDRMKRYNEITAQ-----HKEQSLKDEKSGSNSSE 94

Qy 156 HAGVALAKPLRTPPEQSCCTIPVQEDSPGAPYVNTPTQPKPLGQLCFKQL-- 213

Db 95 SNGIT--PMGTSEKPK-----LQSRTPPSSCIVRHDVTVPKDKN---QGHAFGRLYV 143

Qy 214 ----GEGRLPALPSELHKLIS-----PLQCLNHWKLLHHHPQDGG-PLPLPT--H 256

Db 144 RLHQGRDLNVKSVNAQYAVITFEKTQVMVPPPLKDI-----DGGIPISPKNR 193

Qy 257 PFPYSRLPFPFPFPLOPKPPLPESFLGKLACVDSQKPLDPHLKLACVDSKPLPGP 316

Db 194 PLAGSRGSSGLHS-----ELMLADVRC-----PHWDETVDVTK----- 230

Qy 317 HLEPSCLSRGAHFKSPVEEYLVHVALQGSVSSQASHLSLAKTWAAGSRSEPSKPTD 376

Db 231 --MKSQMVSVYDYEDDKFL-----GSVKITPIFLHEYVQEA-----YKLEPLDLTKS 278

Qy 377 NEGVLLETKLPDYEVREYVHVAHQ-----LRL-GRGSFGVHVMEDKQTFQCAVKVR 432

Db 279 LEGBIKVETI-----YEHIEHRYRGPEDFTALRLIGKGTFGQVYLVRKNDTNRIYAMKKIS 334

Qy 433 LE-VFRAEELMACAGL-----TSPIRVPLYGAVREGFWNI FMELEGSGSLGOLV 481

Db 335 KKLIVRKKEVTHTLGERNILVRTSLDESPFTVGLKFSQFASDLYLITDYMGGSLFWHL 394

Qy 482 KEQCLPEDRALYYLGOALEGLYHRSRIUHGVDKADNVLLSSDGSAAALCDFGHAVCL 541

Db 395 QHEGRFPEQRAKFVIAELVLALHKLHDIYRDLPENILLDADG-HIALCDFG---L 449

Qy 542 QPDGLGKSLTGDYIPGTETHMAPEVVL-GRSCDAKVDMVSSCCMMLHMLGCHPWTQFF 600

Db 450 SKANLSANATNTTC-GTTEVLAPEVLEDDKGYTKQVDFWSLGLVLFEM---CCGWSFFY 505

Qy 601 RGPLCLKIASEPPPVREI-----PSCAPLTAQAIQEG-LRKEPIHRSVAEELG 648

Db 506 -----APDVQOQMYRIAFKRVFRGVLSSSEGRSVRGILNPNHR-----LG 549

Qy 649 GKVNRAQQVGLKSPWRGEYKEPRHPHPNPQANYHTLHAQPRELSPRPGPRPAEETG 708

Db 550 AVADTT-----ELKE--HPFFADINWDL-----LSKK----- 574

Qy 709 RAPKLQPLPPEPPPEPKSPPLTSL---KE--ESGMWEPLPLSSLEPAPARNPSSPERKA 763

Db 575 ---KVQPPF-----KPNVQNDLDVSNFDKFTNTNKNINIVSNVDPANASTPLS----- 621

Qy 764 TVPEQELQQLLELFLNSLSQPF-----SLEEQQLSLCLSIDSLSDSDSEKNPSK 815

Db 622 -----NTIQDRFRGFTFVNKSIDEQFQ-----NLGLQENETDNLH 657

Qy 816 ASQSRDTLSSGVHSSWSQAEARSSWN 843

Db 658 A---CRTTHSSVNSINSNGNPRTVDAN 682

RESULT 13

S58682

protein kinase, p21-activated (EC 2.7.1.1) - human

N;Alternate names: protein kinase PAK65; S6/H4 kinase

C;Species: Homo sapiens (man)

C;Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 16-Aug-2004

C;Accession: S58682; S55258; S55304; S58690; A57441

R;Sells, M.; Knause, U.J.; Bagrodia, S.; Ambrose, D.; Bokoch, G.M.; Chernoff, J.

submitted to the EMBL Data Library, April 1995

A;Description: Human p21-activated protein kinases regulate actin organization in mammal

A;Reference number: S58682

A;Accession: S58682

A;Molecule type: DNA

A;Residues: 1-525 <SEL>

A;Cross-references: UNIPROT:Q13177; EMBL:U24153; NID:g780807; PIDN:AAA65442.1; PID:g7808

R;Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.

EMBO J. 14, 1970-1978, 1995

A;Title: A novel serine kinase activated by rac1/CDC42Hs-dependent autophosphorylation i

A;Reference number: S55258; MUID:95262637; PMID:7744004

A;Accession: S55258

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 'MEETQCKSNLEL', 31-89, 'T', 91-149, 'F', 151-224, 'T', 226-328, 'R', 330-338, 340-525

A;Accession: S55304

A;Molecule type: protein

A;Residues: 402-418 <MAW>

R;Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.

EMBO J. 14, 4385, 1995

A;Reference number: S58690; MUID:96016211; PMID:7556080

A;Contents: erratum

A;Accession: S58690

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-30 <MAP>

R;Benner, G.E.; Dennis, P.B.; Masaracchia, R.A.

J. Biol. Chem. 270, 21121-21128, 1995

A;Title: Activation of an S6/H4 kinase (PAK 65) from human placenta by intramolecular an

A;Reference number: A57441; MUID:95403344; PMID:7673144

A;Accession: A57441

A;Molecule type: protein

A;Residues: 197-216;402,'S',404-409 <BEN>  
A;Experimental source: placenta  
C;Superfamily: serine/threonine protein kinase  
C;Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonine  
F;247-501/Domain: protein kinase homology <KIN>  
F;255-263/Region: protein kinase ATP-binding motif  
F;197,402/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 6.0%; Score 305.5; DB 2; Length 525;  
Best Local Similarity 28.1%; Pred. No. 4.9e-07;  
Matches 94; Conservative 54; Mismatches 160; Indels 27; Gaps 8;

Qy 320 PSLSRGAHFKFSVEYLVALHQLGSSVSSQAHSLTSLAKTWAARGSRSEPSKPTEDNEG 379  
Db 181 PVIAIRPDHKTISYTRSDVDPAPVPGSHVDG-----AAKSLDKQKKPKMTDEE- 231  
Qy 380 VLLTEKLKPDYVEVEEVEVNHATHQLRGSGFGEVHRMEDKQTGQCAVKKVRLVFEFA 439  
Db 232 --IMEKLRTI-VSIGDPKKKYTRYEKIQGASGTFTATDVALGQEAIVKQINLQKQPKK 288  
Qy 440 EL-----MACAGLTSPRIPLYGAVREGPWNIWFELLEGSLGOLVKEQGLPEDRALY 494  
Db 289 ELINIEILLMELKMPNINFLDSYLVGDELPVVMVEYLAGSLTDVVTETACMDAQIAA 348  
Qy 495 YLGOALEGLEYLHSRIILHGDVKADNVLLSSDGSHAAALCDFGHAVCLQPDGIGKSLTGD 554  
Db 349 VCRECLQALEFLHANQVHRIKIDSNVLLMEGS-VKLTDFGCAQIITPEQSKRSTMV-- 405  
Qy 555 YIPGTETHMABEVVLGRSDAKVDVWSSCCMWLHMLNGCHPWQTFRRPL--CLKIASPE 612  
Db 406 ---GTFYMAPEVVRKAYGPKVYDWSLGIAMENVEGEPPYLN--ENPLRALYLIATNG 460  
Qy 613 PPVRIEIPSCAPLTAQATQEGRLKEPIHRVSAEL 647  
Db 461 TPELQNPKEKLSIFRDLNRCLEMDVVERKGSAKEL 495

RESULT 14  
T36746  
probable serine/threonine protein kinase - Streptomyces coelicolor  
C;Species: Streptomyces coelicolor  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: T36746  
R;Saunders, D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1999  
A;Reference number: Z21613  
A;Accession: T36746  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-550 <SAU>  
A;Cross-references: UNIPROT:Q9S2A6; EMBL:AL096849; PIDN:CAB50939.1; GSPDB:GN00070; SCOE  
A;Experimental source: strain A3(2)  
C;Superfamily: serine/threonine protein kinase, PksC type; protein kinase homology

Query Match 6.0%; Score 305; DB 2; Length 550;  
Best Local Similarity 31.7%; Pred. No. 5.4e-07;  
Matches 106; Conservative 37; Mismatches 135; Indels 56; Gaps 13;

Qy 406 LGRGSFGEVHRMEDKQTGQCAVKKVRL-----EVFR-----AEELMACAGLTSPRIV 453  
Db 21 LLEGGMGTWRARDEVLRREVAVKEVRAPAGLSQPDVGRMYARLEREAARISHPNVV 80  
Qy 454 PLYGAVREG--PWNIWFELLEGSLGOLVKEQGLPEDRALYVLGOALEGLEYLHSRI 511  
Db 81 TVYDVATDGGRPW--IVMELVRLSLADILLDAEGPLPRRAALICAEVLAAALRAAAGV 138  
Qy 512 LHGDVKADNVLLSSDGSHAAALCDFGHAVCLQPDGIGKSLTGDYIPGTETHMABEVVLGR 571  
Db 139 LHRDVKPANVLANDG-RVLLTDFGIA---RVESGSEALTMTEGVV-GSPEFLAPRALGR 193  
Qy 572 SCDAKVDVWSSCCMWLHMLNGCHPWQTFRRGLCLKIASPEPPVREIPSCAPLTAQATQ 631

Db 194 TPGAASDLWSLGLVLYATVEGVSPFRQGTPLTLRAIVDEAVP-----PPRRAGALGPVVE 249  
Qy 632 EGIKPEPIHRVSAELGGKVNRAALQOQVG-GLKSPWGEYKEPRHPPPNQANVHOTLHAQP 690  
Db 250 GLLRKDPFAERLPABE-----ABRALRLVGAGAPPGRG-----PRTGAPPSGAFAPITVA-- 299  
Qy 691 RELSPRAGPRPABETTGRAPKLOPPLPPEPPEP 724  
Db 300 ----AHFG-----PPTAPTTPMP 313

RESULT 15  
G96575  
probable MEK kinase MAP3Ka, 84794-81452 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: G96575  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luxos, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A6141; MUID:21016719; PMID:11130712  
A;Accession: G96575  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-608 <STO>  
A;Cross-references: UNIPROT:Q9LPH2; GB:AE005173; NID:g10645340; PIDN:AAG21460.1; GSPDB:  
C;Genetics:  
A;Gene: F22G10.18  
A;Map position: 1

Query Match 6.0%; Score 304; DB 2; Length 608;  
Best Local Similarity 26.4%; Pred. No. 6.6e-07;  
Matches 124; Conservative 67; Mismatches 194; Indels 84; Gaps 21;

Qy 346 SSSQAHSL---TSLAKTWAARGSR---SREPSKPTEDNEGVLTEKLKPDYVEYREHV 398  
Db 175 SSSCHPLPRPTSPTSVAHSGRIGGYETSP-----SGFSTWKKG----- 218  
Qy 399 WATHQLRGRSGFGEVHRMEDKQTGQCAVKKVRL-----EVFRA--BELMACAGLT 448  
Db 219 -----FLGSGTFGQVILGFNSKGMCAIKVKKVSDQTSKECLKQLNQENILNLQIC 272  
Qy 449 SPRIVPLYGAVREGPWNIWFELLEGSLGOLVKEQGLPEDRALYVLGOALEGLEYLHS 508  
Db 273 HPNIQVYVGSSELSEETLSVILEYVSGSIHKLKLDKDYGSFTFVQNYTQRLAGLAYLHG 332  
Qy 509 RRIHLGDKADNVLLSSDGSHAAALCDFG---HAVCLQPDGIGKSLTGDYIPGTETHMAP 565  
Db 333 RNTVHRDIKGANILVDPNG-EIKLADFQWAKHVTAF-----STMLS---FKGSPYVMAP 382  
Qy 566 EVLGRS-CDAKVDVWSSCCMWLHMLNGCHPWQTFRRGLCLKIA-SEPPPVREIPSPCA 623  
Db 383 EVWMSQNGYTHAVDIWSLGTCTILEMATSKPPWQSEGVAAIFKIGNSKDTG---ELPDHLS 440  
Qy 624 PLTAQAIQEGRLKEPIHRVSAEL-LGGKVNRAALQOQVGLKSP-----WRGEYKEP-RH 674  
Db 441 NDAKNFIKLCQRNPTVPTASQLLEHFLRNTTRVASTSLPKDFPPRSYDGNFSLPTRE 500  
Qy 675 PPPQANVHOTLHAQPRELSPRAGPR-----PAEETTGRAPKLOPPL-----PPE 720  
Db 501 PYPGRLS-HDNVAKQPLSRTIKSPRENVRATISLPVSPCSPLRQLGPAYKSCFLSPH 559  
Qy 721 PPE--PNKSPPLTUSKEESGWPEPLPSLEPAPAR--NPSPERKATV 765  
Db 560 PSYAPFGQDSGYNLAEFAASPRMKDAMMEFSSFRQTQTFNPSPLSRVL 608

Search completed: June 7, 2005, 12:03:33  
Job time : 37.0549 secs

---

**THIS PAGE BLANK (USPTO)**